



Quelques contributions à la statistique des Processus : Inégalités de Déviations & Grandes Déviations

Hacène Djellout

► To cite this version:

Hacène Djellout. Quelques contributions à la statistique des Processus : Inégalités de Déviations & Grandes Déviations. Probabilités [math.PR]. Université Blaise Pascal, 2015. tel-01234238

HAL Id: tel-01234238

<https://hal.science/tel-01234238>

Submitted on 26 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ BLAISE PASCAL
(UFR Sciences et Technologies)

Laboratoire de Mathématiques – UMR 6620

Habilitation à Diriger des Recherches

présentée par
Hacène Djellout

Tome I – Synthèse des Travaux

Sujet de l'Habilitation :

Quelques Contributions à la Statistique des Processus :
Inégalités de Déviations & Grandes Déviations

Soutenue le 17 novembre 2015 devant le jury composé de :

Président	Christian LÉONARD, Université Paris Ouest (Paris, France)
Rapporteurs	Fabrice GAMBOA, Université de Toulouse (Toulouse, France) Marc HOFFMANN, Université Paris-Dauphine (Paris, France) Christian HOUDRÉ, Georgia Institute of Technology (Atlanta, USA)
Garant	Arnaud GUILLIN, Université Blaise Pascal (Clermont-Ferrand, France)
Examineurs	Bernard BERCU, Université Bordeaux 1 (Bordeaux, France) Ivan GENTIL, Université Claude Bernard, Lyon 1 (Lyon, France) Li-Ming WU, Université Blaise Pascal (Clermont-Ferrand, France)

*À Kamel disparu en pleine force de l'âge
ton souvenir restera à jamais gravé dans ma mémoire*

Remerciements

Ce mémoire d'habilitation à diriger les recherches doit beaucoup à Arnaud Guillin. Je lui exprime ma plus profonde gratitude pour sa grande disponibilité, ses encouragements constants et pour toutes les discussions et les collaborations que nous avons pu avoir. Arnaud a été d'une aide précieuse. J'admire sa modestie et sa très grande compétence. Le fait d'avoir accepté d'être garant de mon HDR me touche particulièrement.

Je remercie Li-Ming Wu. Après avoir dirigé ma thèse, il a continué à me prodiguer constamment ses conseils et ses encouragements. Li-Ming est une source intarissable d'idées et de questions. J'éprouve un grand respect pour son travail ainsi que pour ses grandes qualités humaines. Je le remercie pour tout ce qu'il m'apprend. Collaborer avec lui est toujours un grand plaisir.

Je suis extrêmement honoré que Fabrice Gamboa, Marc Hoffmann et Christian Houdré aient accepté d'expertiser mon mémoire d'habilitation. Je tiens à les remercier pour avoir consacré une partie de leurs temps précieux à la rédaction des rapports et pour leur présence dans ce jury. Je veux dire aussi merci à Bernard Bercu, Ivan Gentil et Christian Léonard qui ont spontanément accepté de faire partie de mon jury de soutenance.

Je remercie mes co-auteurs Valère Bitseki Penda, Hui Jiang, Arnaud Guillin, Frédéric Proïa et Li-Ming Wu. Je remercie Hui pour son invitation et son accueil très chaleureux à Nanjing en Chine à quelques jours de ma soutenance (Xie Xie!). J'espère que ces collaborations se prolongeront le plus longtemps possible.

Je suis redevable à Thierry Buffard et Erwan Saint Loubert Bié pour la relecture et la correction de la première version du manuscrit.

Je souhaite souligner les très bonnes conditions de travail au sein du laboratoire de mathématiques de l'université Blaise Pascal. J'y apprécie l'ambiance sympathique qui règne entre collègues, et ce quelque soit la spécialité de chacun. Un grand merci aux collègues qui ont partagé avec moi : l'aventure de la coopération avec l'université de Djibouti ou les diverses responsabilités administratives et pédagogiques au sein de l'université notamment celle de la licence MASS.

Je remercie également les organisateurs de l'école d'été de Probabilité de Saint-Flour et à leur tête Laurent Serlet pour leur soutien.

Mes remerciements vont tout particulièrement à ma femme Zhor pour son amour, sa patience et son soutien dont elle fait preuve quotidiennement. Une tendre pensée pour mes deux garçons Harris et Niels, qui m'aident par les joies nombreuses qu'ils m'apportent au quotidien.

Je remercie profondément toute ma famille pour son inconditionnel soutien. Merci pour : ma grand-mère, mes parents, mes frères et soeurs pour leur encouragement et pour la confiance et le respect qu'ils m'accordent.

Table des matières

Bibliographie personnelle	3
Introduction	5
1 Grandes déviations pour les estimateurs de la (co-)volatilité	11
1.1 Introduction	11
1.2 Grandes déviations dans le cas synchrone et sans sauts	12
1.3 Grandes déviations dans le cas synchrone avec sauts	18
1.4 Grandes déviations dans le cas non-synchrone sans saut	19
Problèmes et perspectives	20
2 Chaînes de Markov bifurcantes, théorèmes limites	23
2.1 Introduction	23
2.2 Théorèmes limites pour des Chaînes de Markov Bifurcantes	25
2.2.1 Cadre des fonctionnelles non-bornées	26
2.2.2 Cadre des fonctionnelles bornées	28
2.2.3 Application : Processus autorégressif bifurcant du premier ordre	29
2.3 Processus auto-régressif bifurcant d'ordre p	32
2.3.1 Principaux résultats	34
Problèmes et perspectives	36
3 Inégalités fonctionnelles et applications	39
3.1 Introduction	39
3.2 Inégalités de coût du transport et applications	40
3.2.1 Principaux résultats sur T_p	40
3.2.2 Inégalité de transport pour des variables aléatoires dépendantes	41
3.2.3 Inégalité de transport pour des systèmes dynamiques aléatoires	43
3.3 Inégalités fonctionnelles par l'estimation en norme Lipschitzienne d'une équation de Poisson	46
3.3.1 Principaux résultats	48
3.3.2 Applications aux inégalités de transport et de concentration	50
3.3.3 Exemples	52
Problèmes et perspectives	54
4 Déviations modérées pour des variables aléatoires dépendantes	55
4.1 Introduction	55
4.2 Chaînes de Markov	56
4.3 Différences de martingales et suites ϕ -mélangeantes	58
4.4 Estimation de la densité spectrale d'un processus à moyenne mobile	61
4.5 Statistique de DURBIN-WATSON	63
Problèmes et perspectives	64
Bibliographie	65

Bibliographie Personnelle¹

Thèse

- [T] H. DJELLOUT, Grandes déviations et déviations modérées de processus stochastique. *Thèse de l'Université Blaise Pascal*, (2000).

Articles dans des Revues avec Comité de Lecture

- [T-D1] H. DJELLOUT, A. GUILLIN and L. WU, Large and moderate deviation for estimators quadratic variationnal processes of diffusions. *Statistical Inference For Stochastic Processes*, 2(3), 195–225, 1999.
- [T-D2] H. DJELLOUT and A. GUILLIN, Moderate deviation for Markov chains with atom. *Stochastic Processes and Their Applications*, 95(2), 237–251, 2001.
- [T-D3] H. DJELLOUT and A. GUILLIN, Large and moderate deviation for moving average processes. *Annales de la Faculté des Sciences de Toulouse*, X(1), 23–31, 2001.
- [T-D4] H. DJELLOUT, Moderate deviations for martingale differences and applications to ϕ -mixing sequences. *Stochastics and Stochastics Reports*, 73(1-2), 37–63, 2002.
- [H-D5] H. DJELLOUT, A. GUILLIN and L. WU, Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Annals of Probability*, 32(3B), 2702–2732, 2004.
- [H-D6] H. DJELLOUT, A. GUILLIN and L. WU, Moderate deviations of empirical periodogram and non-linear functionals of moving average processes. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 42(4), 393–416, 2006.
- [H-D7] H. DJELLOUT and L. WU, Lipschitzian norm estimate of one dimensional Poisson equations. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 47, no. 2, 450–465, 2011.
- [H-D8] V. BITSEKI PENDA, H. DJELLOUT and A. GUILLIN, Deviation inequalities, Moderate deviations and some limit theorems for bifurcating Markov chains with application. *Annals of Applied Probability*, 24, no. 1, 235–291, 2014.
- [H-D9] V. BITSEKI PENDA, H. DJELLOUT and F. PROIA, A moderate deviation principle for the Durbin-Watson statistic associated with the stable first-order autoregressive process. *Accepté pour publication dans ESAIM : Probability and Statistics*. (2014).
- [H-D10] V. BITSEKI PENDA and H. DJELLOUT, Deviation inequalities and moderate deviations for estimators of parameters in bifurcating autoregressive models. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 50, no. 3, 806–844, 2014.
- [H-D11] V. BITSEKI PENDA, H. DJELLOUT, L. DUMAZ, F. MERLEVEDE, F. PROIA, Moderate deviations of functional of Markov Processes. *Accepté pour publication dans ESAIM : Proceedings*. (2014).
- [H*-D12] H. DJELLOUT and Y. SAMOURA, Large and moderate deviations of realized covolatility. *Statistics & Probability Letters*, 86, 30–37, 2014.

1. T : publication pendant la thèse, H : publication après la thèse, H* : publication avec le doctorant.

Articles soumis à publication

- [H*-D13] H. DJELLOUT, A. GUILLIN and Y. SAMOURA, Large deviations of the realized (co-) volatility vector. Soumis 2014.
- [H-D14] H. DJELLOUT and H. JIANG, Large deviations of the Threshold estimator of integrated (co-) volatility vector in the presence of jumps. Soumis 2015.

Articles en en préparation

- [H*-D15] H. DJELLOUT, H. JIANG, A. GUILLIN and Y. SAMOURA, Moderate deviations for the (co-)volatility vector in the non-synchronous observations case.
- [H-D16] H. DJELLOUT, A. GUILLIN and L. WU, Large deviations for symmetric Markov Chain with applications to estimations of the volatility.

Notes aux Comptes Rendus de l'Académie des Sciences

- [T-D17] H. DJELLOUT et A. GUILLIN, Principe de déviations modérées pour le processus empirique fonctionnel d'une chaîne de Markov. *Comptes rendus de l'Académie des Sciences- Série I, tome 330, Mathématiques*, 377-380. 2000.

Rapports de Recherches

- [T-D18] H. DJELLOUT, Sur l'unicité de l'opérateur de Schrödinger unidimensionnel, Mémoire de DEA 1997.
- [T-D19] H. DJELLOUT, A. GUILLIN and L. WU, Large and moderate deviations for functionals depending of infinite variables of an i.i.d. sequence. 2001.

Introduction

Mes principales activités de recherche concernent l'étude des grandes déviations, des déviations modérées et des inégalités de déviations pour certains processus aléatoires issus de la statistique (estimateurs des paramètres d'une diffusion, chaînes de Markov, martingales, moyennes mobiles et périodogramme). J'ai diversifié mes activités de recherches en m'intéressant aux inégalités de transport-information, à la statistique des processus, aux chaînes de Markov bifurcantes et aux modèles de vieillissement de population, ainsi qu'aux mathématiques financières.

Ce document présente une synthèse de mes travaux de recherche et décrit ainsi brièvement les articles dont la liste est présentée p.3. Ces travaux incluent tous les résultats obtenus depuis ma thèse jusqu'aux plus récents. Ces recherches se regroupent naturellement selon 4 thèmes principaux, ne respectant pas l'ordre chronologique, mais qui ont en commun une thématique majeure : les **grandes déviations**, et par extension, les **déviations modérées**, les **inégalités de déviation** et leurs applications. Ces quatre thèmes sont les suivants :

1. Grandes déviations pour les estimateurs de la (co)-volatilité

Ces travaux sont faits après la thèse, sauf le premier qui a été réalisé pendant celle-ci.

Les processus de diffusion jouent maintenant un rôle prédominant dans de nombreux champs applicatifs comme les mathématiques financières ou la génétique des populations. Il est alors crucial de pouvoir estimer finement le coefficient de diffusion (appelé *volatilité* en finance), associé à ces équations différentielles stochastiques.

Dans l'estimation de la volatilité en finance, on dispose souvent d'observations à des temps, régulièrement espacés ou non, ou encore à des temps aléatoires, ce qui rend le problème plus difficile. Depuis quelques années, des percées méthodologiques ont permis de montrer la consistance ou la normalité asymptotique des estimateurs, dans ces cas. Cependant, afin de calibrer finement les modèles, il est crucial de pouvoir montrer des théorèmes plus puissants de type grandes déviations, c'est-à-dire d'établir des équivalents asymptotiques logarithmiques pour l'erreur d'estimation. C'est un problème largement ouvert, quelques résultats seulement existent qui sont souvent restreints à des cas d'observations uniformes.

Plus précisément, nous modélisons l'évolution d'une variable d'état observable par un processus stochastique $X_t = (X_{1,t}, X_{2,t})$, $t \in [0, 1]$. Dans les applications financières, X_t peut être considéré comme le taux d'intérêt à court terme, un taux de change ou le logarithme du prix d'un actif ou d'un indice boursier. Supposons que $X_{1,t}$ et $X_{2,t}$ sont définies sur un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ et suivent un processus de Itô, à savoir :

$$\begin{cases} dX_{1,t} &= \sigma_{1,t}dB_{1,t} + b_1(t, \omega)dt, \\ dX_{2,t} &= \sigma_{2,t}dB_{2,t} + b_2(t, \omega)dt, \end{cases}$$

où B_1 et B_2 sont des mouvements browniens standards corrélés avec $\text{Corr}(B_{1,t}, B_{2,t}) = \rho_t$. On peut écrire $dB_{2,t} = \rho_t dB_{1,t} + \sqrt{1 - \rho_t^2} dB_{3,t}$, où $B_1 = (B_{1,t})_{t \in [0,1]}$ et $B_3 = (B_{3,t})_{t \in [0,1]}$ sont des mouvements browniens indépendants.

On limite notre attention au cas où σ_1 , σ_2 et ρ sont des fonctions déterministes. Les fonctions σ_ℓ , $\ell = 1, 2$ prennent des valeurs positives tandis que ρ prend ses valeurs dans l'intervalle $] -1, 1[$.

Nous nous sommes intéressés au vecteur de la (co)-variation

$$V_t = ([X_1]_t, [X_2]_t, \langle X_1, X_2 \rangle_t)^T$$

où $[X_\ell]_t := \int_0^t \sigma_{\ell,s}^2 ds$, $\ell = 1, 2$ représente la variation quadratique du processus X_ℓ , et $\langle X_1, X_2 \rangle_t := \int_0^t \sigma_{1,s} \sigma_{2,s} \rho_s ds$ la covariance déterministe entre X_1 et X_2 .

Étant donné des observations discrètes également espacées $(X_{1,t_k^n}, X_{2,t_k^n}, k = 1, \dots, n)$ dans l'intervalle $[0, 1]$ (avec $t_k^n = k/n$), un théorème limite pour les processus stochastiques affirme que

$$(1) \quad V_t^n(X) = (Q_{1,t}^n(X), Q_{2,t}^n(X), C_t^n(X))^T,$$

usuellement appelée la (co-)variance réalisée, est un estimateur consistant de V_t , avec, pour $\ell = 1, 2$

$$Q_{\ell,t}^n(X) := \sum_{k=1}^{[nt]} \Delta_k X_\ell^2 \quad \text{et} \quad C_t^n(X) := \sum_{k=1}^{[nt]} \Delta_k X_1 \Delta_k X_2 \quad \text{où} \quad \Delta_k X_\ell := X_{\ell,t_k} - X_{\ell,t_{k-1}}.$$

Notre but dans ce chapitre est de présenter les principes de grandes déviations et de déviations modérées de cet estimateur. Bien que ces statistiques aient été étudiées depuis près de 20 ans, il y a de façon surprenante très peu de résultats dans cette direction. L'explication peut être la suivante : les techniques habituelles (telle que la méthode de Gärtner-Ellis) ne fonctionnent pas et un traitement très particulier doit être effectué pour ce problème.

Notons que la forme des fonctions de taux de grandes déviations est aussi originale. En effet, un terme correctif supplémentaire apparaît dans l'évaluation de la fonction de taux, au niveau processus. Ce phénomène est dû à la faible intégrabilité exponentielle des estimateurs, il a été d'abord remarqué par Lynche et Sethuraman [119].

Nous avons considéré le cas unidimensionnel dans [T-D1], un travail en collaboration avec A. GUILLIN et L. WU [T-D1] publié dans *Statistical Inference For Stochastic Processes*. Nous avons étudié les grandes déviations et les inégalités de déviations pour l'estimateur de la variation quadratique $Q_{1,t}^n(X)$.

Dans le cas bidimensionnel, nous avons obtenu en collaboration avec Y. SAMOURA [H-D12], publié dans *Statistics & Probability Letters*, le principe de grandes déviations pour l'estimateur de la covolatilité réalisée $C_t^n(X)$.

Récemment, en collaboration avec A. GUILLIN et Y. SAMOURA ([H*-D13] soumis), nous avons considéré les grandes déviations pour le vecteur $V_t^n(X)$ défini dans (1). Comme application de notre résultat principal, nous avons donné, dans certains cas, les grandes déviations pour les estimateurs des coefficients de corrélation et de régression.

En collaboration avec H. JIANG ([H-D14] soumis), nous considérons une diffusion bidimensionnelle en présence de processus de sauts de type processus de Poisson composé de la forme $J_{\ell,t} = \sum_{i=1}^{N_{\ell,t}} Y_{\ell,i}$. Les processus à sauts sont des outils standards, par exemple, pour modéliser des valeurs d'un actif financier ou dans l'assurance. La motivation clé derrière des modèles de diffusion à sauts est l'étude de marché des actions qui aboutissent à des changements soudains du prix ou l'étude des titres risqués et qui ne peuvent pas être modélisés par une seule composante diffusive continue.

L'estimateur $V_t^n(X)$ donné dans (1) est fortement influencé quand les processus X_ℓ contiennent des sauts. En fait, quand n tend vers l'infini une telle somme s'approche du vecteur (co-)variation quadratique globale qui contient aussi les co-sauts.

Au cours des dernières années, plusieurs auteurs ont proposé diverses méthodes d'estimation pour le vecteur de la (co-)volatilité. L'utilisation d'un estimateur tronqué est une des techniques utilisées dans l'inférence pour des processus stochastiques à sauts. Dans ce travail, nous utilisons l'estimateur tronqué introduit par Mancini [121] où seulement les variations par rapport à une fonction de seuil donnée sont prises en compte. On utilise donc l'estimateur $\mathcal{V}_t^n = (Q_{1,t}^n, Q_{2,t}^n, C_t^n)$ où :

$$\mathcal{Q}_{\ell,t}^n = \sum_{k=1}^{[nt]} (\Delta_k X_\ell)^2 \mathbf{1}_{\{(\Delta_k X_\ell)^2 \leq r(\frac{1}{n})\}}, \quad \text{et} \quad C_t^n = \sum_{k=1}^{[nt]} \Delta_k X_1 \Delta_k X_2 \mathbf{1}_{\{\max_{\ell=1}^2 (\Delta_k X_\ell)^2 \leq r(\frac{1}{n})\}}.$$

Sous certaines conditions sur la fonction de seuil, nous obtenons les mêmes grandes déviations que dans le cas de processus sans saut.

Actuellement, en collaboration avec H. JIANG, A. GUILLIN et Y. SAMOURA [H-D15], nous considérons une diffusion bidimensionnelle dans le cas d'un échantillonnage asynchrone, en utilisant l'estimateur de Hayashi-Yoshida défini par

$$U_{n,m} := \sum_{i=1}^n \sum_{j=1}^m \Delta X_1(I^i) \Delta X_2(J^j) I_{\{I^i \cap J^j \neq \emptyset\}},$$

où $I^i = (s_{i-1}, s_i]$, $J^j = (t_{j-1}, t_j]$. Hayashi et Yoshida ont prouvé que cet estimateur était à la fois non biaisé pour la covariance et consistant. On obtient les déviations modérées dans ce cas. Les grandes déviations sont un peu plus difficiles à obtenir, sauf dans quelques cas particuliers pour les coefficients de diffusion.

2. Chaînes de Markov bifurcantes et théorèmes limites

Ces travaux ont été effectués après la thèse.

Les chaînes de Markov bifurcantes (BMC) sont une adaptation des chaînes de Markov habituelles aux données d'un arbre binaire régulier. En d'autres termes, ce sont des chaînes de Markov pour lesquelles l'ensemble des indices est un arbre binaire régulier. Elles sont appropriées par exemple dans la modélisation de données d'origine cellulaires quand chaque cellule d'une génération donne naissance à deux descendants dans la génération suivante.

Récemment, les chaînes de Markov bifurcantes ont fait l'objet de beaucoup de travaux à cause d'expériences de biologistes étudiant le vieillissement de *Escherichia Coli* (voir [160], [97]). *E. Coli* est une bactérie en forme de tige qui se reproduit en se divisant en son milieu, produisant ainsi deux cellules : celle qui a déjà existé et que nous appelons la vieille descendante du pôle, et l'autre qui est nouvelle et que nous appelons la nouvelle descendante du pôle. Le but de ces expériences était de chercher la trace de vieillissement dans *E. Coli*.

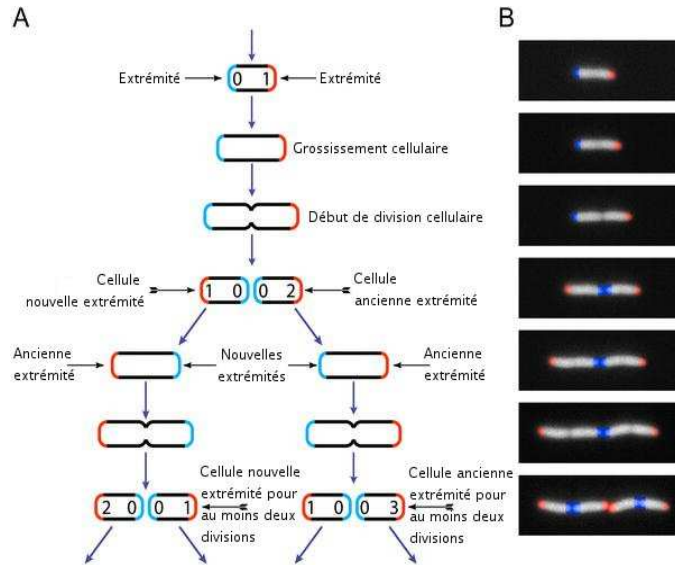


FIGURE 1 – Mécanisme de division, de Stewart, Madden, Paul et Taddei, *PLoS Biol.*, 2005

Cowan et Staudte [50] et Guyon & al. ([97], [96]) ont proposé le modèle Gaussien linéaire suivant pour décrire l'évolution du taux de croissance X_n de l'individu n de la population des cellules tirées d'un

individu initial

$$\mathcal{L}(X_1) = \nu, \quad \text{et} \quad \forall n \geq 1, \quad \begin{cases} X_{2n} = \alpha_0 X_n + \beta_0 + \varepsilon_{2n} \\ X_{2n+1} = \alpha_1 X_n + \beta_1 + \varepsilon_{2n+1}, \end{cases}$$

n étant la mère, $2n$ la nouvelle descendante du pôle et $2n+1$ la vieille descendante du pôle, $\alpha_0, \alpha_1 \in (-1, 1)$; $\beta_0, \beta_1 \in \mathbb{R}$ et $(\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1$ des variables aléatoires i.i.d..

Ce modèle est un exemple typique de la dynamique Markovienne bifurcante et a été la motivation pour l'étude mathématique rigoureuse des BMC [96].

On note \mathbb{T}_r le sous-arbre de tous les individus en partant de l'individu initial jusqu'à la r -ième génération. On sait que son cardinal $|\mathbb{T}_r|$ est égal à $2^{r+1} - 1$. Nous nous sommes intéressés aux moyennes empiriques du type

$$M_{\mathbb{T}_r}(f) = \frac{1}{|\mathbb{T}_r|} \sum_{i \in \mathbb{T}_r} f(X_i, X_{2i}, X_{2i+1}).$$

Un outil essentiel est la chaîne induite $Y = (Y_i, i \in \mathbb{N}^*)$ correspondant à une lignée prise uniformément au hasard parmi toutes les lignées.

Nos objectifs dans ce travail en collaboration avec V. BITSEKI PENDA et A. GUILLIN ([H-D8] publié dans *Annals of Applied Probability*), sont multiples. Sous l'hypothèse d'ergodicité géométrique ou d'ergodicité géométrique uniforme de cette chaîne, on obtient quelques théorèmes limites comme la loi des grands nombres, la loi du logarithme itéré, le théorème de la limite centrale et le principe de déviations modérées pour une classe de fonctions f . Nous obtenons aussi des inégalités de déviations pour la quantité précédente. Nos résultats mettent en évidence une compétition entre la division binaire et la vitesse de convergence du coefficient d'ergodicité et on obtient des régimes dépendant de la position de ce coefficient par rapport à $1/2$.

Comme applications de ces résultats sur les chaînes de Markov bifurcantes, on obtient les déviations modérées et les inégalités de déviations pour l'estimateur des paramètres inconnus dans le processus autorégressif bifurcant d'ordre 1 avec un bruit Gaussien ou un bruit borné. Ces estimations sont importantes pour une étude statistique rigoureuse. En collaboration avec V. BITSEKI PENDA ([H-D8] publié dans *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*), on a étudié les mêmes questions mais pour un modèle auto-régressif d'ordre quelconque par une approche martingale. Une analyse fine de ce modèle est faite par BERCU, GÉGOUT-PETIT et SAPORTA (2009).

3. Inégalités fonctionnelles et applications

Ces travaux ont été faits après la thèse.

Nous nous intéressons dans cette partie aux inégalités fonctionnelles du type inégalités de transport ou de Sobolev logarithmiques, dont l'étude a suscité ces 20 dernières années un engouement croissant, voir par exemple l'ouvrage pédagogique [6] et les cours, ou le livre de Ledoux [114, 115]. Leur intérêt provient non seulement de leurs conséquences en probabilités en termes, par exemple, d'inégalités de déviations, de concentration de la mesure et de convergence de semi-groupes, mais aussi en analyse des équations aux dérivées partielles, pour l'étude de la vitesse d'atteinte de l'équilibre, voir par exemple Villani [167].

Sur (E, d) un espace métrique, pour $p \geq 1$ et μ, ν deux mesures de probabilités sur E , on définit la distance de Wasserstein L^p

$$W_p^d(\mu, \nu) = \inf \left(\int \int d(x, y)^p d\pi(x, y) \right)^{1/p}$$

où l'infimum est pris sur toutes les mesures de probabilité sur $E \times E$ de marginales μ et ν (π est alors appelée couplage de μ et ν). On dit alors que μ satisfait une inégalité de transport T_p , pour $p \in [1, 2]$, de constante C (que nous notons $\mu \in T_p(C)$), si pour toute mesure de probabilité ν

$$W_p^d(\mu, \nu) \leq \sqrt{2CH(\nu|\mu)}$$

avec $H(\nu|\mu)$ l'entropie (ou information de Kullback) définie par

$$H(\nu|\mu) = \int \log \frac{d\nu}{d\mu} d\nu, \quad \text{si } d\nu \ll d\mu; +\infty \text{ sinon.}$$

Ces inégalités impliquent, de par les travaux de MARTON et TALAGRAND, des inégalités de concentration de type gaussienne et sont donc particulièrement utiles en pratique. Rappelons que l'inégalité de concentration se traduit par une des deux formulations équivalentes suivantes :

- pour $A \subset E$ avec $\mu(A) \geq 1/2$ (par exemple), alors $\mu(A^r) \geq 1 - \varepsilon(r)$ où $A^r = \{x \in E, d(x, A) \leq r\}$,
 - pour toute fonction f 1-Lipschitz (par exemple), $\mu(\{x \in E, f(x) \geq m + r\}) \leq \varepsilon(r)$, m étant la médiane.
- On parle d'une concentration Gaussienne quand $\varepsilon(r) \simeq e^{-cr^2}$.

Notons que pour $p = 1$ et $d(x, y) = 1_{x \neq y}$, T_1 est l'inégalité de Pinsker-Csizsar, qui est toujours vérifiée. Pour des distances générales, les inégalités de transport ont été étudiées par Marton [124, 125] qui les a élégamment reliées au phénomène de concentration de la mesure; par Talagrand [163] qui a montré que T_2 était vérifiée par la mesure gaussienne et la distance euclidienne; par BOBKOV-GÖTZE [33] qui ont donné une caractérisation de T_1 par un contrôle de la transformée de Laplace associée à la mesure, (voir Ledoux [115] pour de plus amples références). Otto-Villani [141] et Bobkov-Gentil-Ledoux [32] ont montré qu'une inégalité de Sobolev logarithmique impliquait T_2 .

En collaboration avec A. GUILLIN et L. WU ([H-D5] publié dans *Annals of Probability*), nous avons cherché une condition nécessaire et suffisante applicable pour le transport T_1 comme pour les mesures satisfaisant une condition d'intégrabilité gaussienne, complétant la caractérisation fonctionnelle établie par Bobkov-Götze. Ainsi μ satisfait une inégalité T_1 si et seulement si :

$$\exists \delta > 0, \quad \int \int e^{\delta d(x, y)^2} d\mu(x) d\mu(y) < \infty.$$

Nous montrons ensuite comment étendre ces inégalités de transport aux suites dépendantes (notamment markoviennes) et nous les appliquons au cas de lois de diffusions ainsi qu'à certains systèmes dynamiques aléatoires. Nous proposons ensuite une méthode de type GIRSANOV pour établir le transport T_2 dans le cas infini-dimensionnel de la mesure de WIENER ou de lois de diffusions, obtenant ainsi des résultats pour ce transport qui ne découlent pas d'une inégalité de Log- SOBOLEV.

Nous obtenons dans un travail en collaboration avec L. WU ([H-D7], publié dans *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*,) des inégalités transport-information plus précises pour la moyenne empirique d'un processus de Markov unidimensionnel en contrôlant la norme Lipschitzienne de la solution de l'équation de Poisson. Plus précisément, considérons $(X_t : t \geq 0)$ une diffusion réversible par rapport à une probabilité μ , de générateur infinitésimal \mathcal{L} . On s'intéresse à la solution $G \in L^2(\mu)$ avec $\mu(G) = 0$ de l'équation de Poisson

$$-\mathcal{L}G = g$$

où $g \in L^2(\mu)$ telle que $\mu(g) = 0$. Dans l'interprétation physique de la diffusion de chaleur, g représente la source de chaleur et G est la distribution de chaleur d'équilibre. Cette inversion du générateur nous donne des informations importantes sur le processus de Markov associé, spécialement en ce qui concerne la moyenne empirique associée. On se restreint au cas unidimensionnel sur un intervalle I . Notre objectif est d'estimer

$$\|G\|_{Lip(\rho)} := \sup_{x, y \in I, x < y} \frac{|G(y) - G(x)|}{|\rho(y) - \rho(x)|}$$

à l'aide des différentes normes de g , (ici ρ est telle que $\rho'(x) > 0$). Ceci conduit à une interprétation intéressante de la formule variationnelle pour le trou spectral λ_1 due à Chen ([43]).

L'importance du trou spectral réside dans le fait qu'il décrit le taux de convergence exponentielle. L'étude de λ_1 a une très longue histoire, le lecteur est renvoyé aux livres [44, 168] pour des bibliographies.

Nous obtenons dans ce travail la meilleure constante dans l'inégalité isopérimétrique de CHEEGER et certaines inégalités de concentration plus précises pour la moyenne empirique

$$\frac{1}{t} \int_0^t g(X_s) ds.$$

Ces résultats sont des conséquences immédiates de l'estimation de $\|G\|_{Lip(\rho)}$ *via* la décomposition forward-backward de martingale ou *via* des inégalités transport-information développées dans [95].

4. Déviations modérées pour des variables aléatoires dépendantes

Dans ces travaux, faits essentiellement pendant ma thèse (5 publications pendant la thèse et 2 après la thèse), on s'intéresse aux principes des grandes déviations et des déviations modérées de fonctionnelles dépendant d'une suite infinie de variables aléatoires indépendantes de même loi. Ce cadre couvre diverses situations : filtrage, statistique, systèmes dynamiques, moyennes mobiles [T-D3] et [T-D19]. Nous nous sommes intéressés à l'obtention des conditions les plus optimales pour le principe des déviations modérées pour des différences de martingales, pour les suites ϕ -mélangeantes [T-D4], pour le périodogramme empirique [H-D6], pour le processus empirique fonctionnel d'une chaîne de Markov [T-D2], [T-D17] et la statistique de Durbin-Watson [H-D9], en s'inspirant du cadre indépendant. Nous renvoyons au dernier chapitre de ce manuscrit pour plus de détails.

Rappel

Par souci de clarté et pour ne pas alourdir ce manuscrit, nous ne présenterons que les éléments clés des démonstrations de nos principaux résultats, oubliant parfois les hypothèses les plus techniques pour ne laisser que les plus significatives, renvoyant aux textes complets des articles pour les détails.

Enfin, comme ce manuscrit se place clairement dans le cadre des grandes déviations, rappelons la définition de celles-ci : soit (X_n) une famille de variables aléatoires définies sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans un espace topologique séparé et régulier E . On dit que (X_n) (ou $\mathbb{P}(X_n \in \cdot)$) satisfait le Principe de Grandes Déviations (PGD) de vitesse λ_n tendant vers l'infini et de bonne fonction de taux I si

1. $I : E \rightarrow [0, \infty]$ est une application semi-continue inférieurement et inf-compacte (ensembles de niveaux compacts) ;
2. pour tout ensemble borélien A de E , on a

$$-\inf_{x \in A^\circ} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{\lambda_n} \log \mathbb{P}(X_n \in A) \leq \limsup_{n \rightarrow \infty} \frac{1}{\lambda_n} \log \mathbb{P}(X_n \in A) \leq -\inf_{x \in \bar{A}} I(x),$$

où A° et \bar{A} désignent respectivement, l'intérieur et la fermeture de A .

On dit que la suite de variables aléatoires (X_n) satisfait le Principe de Déviations Modérées (PDM) de vitesse b_n^2 avec $1 \ll b_n \ll \sqrt{n}$ et de fonction de taux J , si la suite $(\sqrt{n}(X_n - \mathbb{E}(X_n))/b_n)_n$ satisfait le PGD de vitesse b_n^2 et de fonction de taux J .

Un PDM est donc un résultat intermédiaire entre le théorème de la limite centrale (TCL) et un PGD. Un PDM nécessite généralement des hypothèses moins fortes qu'un PGD et présente une fonction de taux généralement quadratique, plus aisément manipulable. Il a en outre un lien naturel avec la loi du logarithme itéré (voir Deuschel-Stroock [66]).

On aura besoin de la notion de convergence super exponentielle, que nous rappelons ici. On dit que X_n converge λ_n -super exponentiellement vers une variable aléatoire X en probabilité, où $\lambda_n \rightarrow \infty$, si pour tout $\delta > 0$

$$\limsup_{n \rightarrow \infty} \frac{1}{\lambda_n} \log \mathbb{P}(d(X_n, X) > \delta) = -\infty.$$

Cette convergence super exponentielle de vitesse λ_n sera notée $X_n \xrightarrow[\lambda_n]{\text{superexp}} X$.

Chapitre 1

Grandes déviations pour les estimateurs de la (co-)volatilité

Dans ce chapitre, nous présentons les articles [T-D1], publié dans *Statistical Inference For Stochastic Processes*, [H-D12], publié dans *Statistics & Probability Letters*, les articles [H-D13] et [H-D14] sont soumis et l'article [H-D15] est un travail en cours de réalisation.

1.1 Introduction

Pendant les dix dernières années, il y a eu un développement considérable de la théorie asymptotique pour des processus observés à une haute fréquence. Ceci a été principalement motivé par des applications financières, où les données, comme les cours des actions ou des monnaies, sont observées très fréquemment.

La covariance de la rentabilité d'un actif et ses statistiques connexes jouent un rôle de premier plan dans de nombreux problèmes théoriques ainsi que pratiques, rencontrés dans le domaine de la finance. De façon analogue à l'approche de la volatilité réalisée, l'idée d'employer des données haute fréquence dans le calcul de la covariance entre deux actifs mène au concept de covariance réalisée. Le rôle clé de la quantification de la (co-)volatilité intégrée dans l'optimisation de portefeuille et de gestion des risques, a stimulé un intérêt croissant pour les méthodes d'estimation pour ces modèles.

Il est naturel d'utiliser le cadre asymptotique lorsque le nombre d'observations à haute fréquence dans un intervalle de temps fixe (un jour par exemple) augmente à l'infini.

Ainsi Barndorff-Nielsen et Shephard [16] ont établi une loi des grands nombres, et étudié les fluctuations correspondantes de la volatilité réalisée. Ces estimations ont été étendues aux configurations plus générales de la statistique par Barndorff-Nielsen et *al.* [15] et [14]. Dovonon, Gonçalves et Meddahi [69] ont considéré des expansions d'Edgeworth pour la statistique de la volatilité réalisée et son analogue bootstrap.

Ces résultats sont essentiels pour explorer les comportements asymptotiques de la (co-)volatilité réalisée. Il y a aussi des estimations par des approches différentes pour la (co-)volatilité intégrée dans les modèles multidimensionnels et les théorèmes limites. On peut se référer à Barndorff-Nielsen et *al.* [17] et [15] où les auteurs présentent, d'une manière unifiée, une loi faible des grands nombres et un TCL pour un estimateur général, appelé variation réalisée généralisée à double puissance.

Pour les travaux connexes concernant les cas de deux variables en vertu d'un plan d'échantillonnage non synchrone, voir Hayashi et Yoshida [99], Bibinger [26], Dalalyan et Yoshida [51], voir aussi Aït-Sahalia et *al.* [2] et les références qui y sont citées. Pour l'estimation de la (co-)variance du processus log-prix en présence de bruit de microstructure du marché, nous renvoyons à Bibinger et Reiss [27], Robert et Rosenbaum [152], Zhang et *al.* [184] et [185]. Voir aussi Gloter [87], ou Comte et *al.* [49] pour l'estimation non paramétrique dans le cas d'un modèle à volatilité stochastique.

Le problème des grandes déviations apparaissant dans la théorie de l'inférence statistique est tout naturel. Pour l'estimation des paramètres et des fonctions inconnues, il est important de minimiser le risque de mauvaises décisions déduites des déviations entre les valeurs observées des estimateurs et les vraies valeurs de paramètres ou des fonctions à estimer. Les grandes déviations des estimateurs peuvent nous fournir des taux de convergence, ainsi qu'une méthode utile pour construire des intervalles de confiance asymptotiques.

Le but de ce chapitre est de présenter les grandes déviations des estimateurs de la (co)-volatilité dans diverses situations. Bien que ces statistiques aient été étudiées depuis 20 ans, il y a remarquablement peu de résultats dans cette direction, ce qui n'a pas manqué de nous surprendre.

L'explication peut toutefois être la suivante : les techniques habituelles (telles que la méthode de Gärtner-Ellis) ne fonctionnent pas, et un traitement très particulier doit être effectué pour ce problème.

Nous avons considéré le cas unidimensionnel dans [T-D1] et récemment Kanaya et Otsu [108] ont obtenu les grandes déviations pour la volatilité réalisée, mais leurs résultats sont déjà contenus dans notre travail [T-D1]. Dans le cas bidimensionnel, nous avons obtenu, en collaboration avec Samoura [H-D12], les grandes déviations pour la covolatilité réalisée. Dans le cas d'une diffusion unidimensionnelle avec une composante à sauts, le principe de grandes déviations pour l'estimateur tronqué de la volatilité constante a été établi par Mancini [122]. Le principe de déviations modérées trajectoriel pour cet estimateur du processus variationnel quadratique a été obtenu par Jiang [107].

Comme on l'a signalé la stratégie de preuve de Gärtner-Ellis des grandes déviations ne peut pas être adaptée dans ces cas. Nous sommes dans la situation que Schied [156] appelle "la condition faible de Cramer". Nous rencontrons les mêmes difficultés techniques que dans les papiers de Bercu et *al.* [23], Bryc et Dembo [40] et Gamboa et *al.* [82] où ils ont établi les grandes déviations pour les formes quadratiques de processus gaussiens. Sur les grandes déviations des variables aléatoires pondérées, nous nous référons aussi à Maïda et *al.* [120] et [182].

Puisque nous ne pouvons pas déterminer le comportement limite de la fonction génératrice des moments à un certain point de la frontière du domaine de définition, nous utilisons d'autres approches. Notons que la forme des fonctions de taux des grandes déviations est aussi originale : au niveau du processus et à cause de la faible intégrabilité exponentielle des estimateurs, un terme correctif supplémentaire apparaît dans l'évaluation de la fonction de taux, phénomène d'abord découvert par Lynche et Sethuraman [119].

Dans nos approches, nous nous sommes principalement inspirés des travaux suivants : les résultats de Lynch et Sethuraman [119] sur les grandes déviations de processus à accroissements indépendants et homogènes, et ceux de Puhalski [145] sur les grandes déviations de processus stochastiques ; les travaux de Bryc et Dembo [40], et Bercu, Gamboa et Rouault [23] sur les grandes déviations de formes quadratiques de processus gaussiens stationnaires. Nous avons utilisé aussi une autre approche basée sur les résultats de Najim [134, 135, 133], où l'hypothèse d'escarpement de la fonction génératrice des moments est relâchée.

Pour les résultats originaux sur les grandes déviations trajectorielles on peut citer les travaux de de-Acosta [52], Borovkov [37], Varadhan [166] et Moguslkii [131], [132].

1.2 Grandes déviations dans le cas synchrone et sans sauts

Nous modélisons l'évolution d'une variable observable d'état par un processus stochastique $X_t = (X_{1,t}, X_{2,t}), t \in [0, 1]$. Dans les applications financières, X_t peut être considéré comme le taux d'intérêt à court terme, un taux de change ou le logarithme du prix d'un actif ou d'un indice boursier.

Supposons que $X_{1,t}$ et $X_{2,t}$ soient définies sur un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfaisant aux conditions usuelles, et suivent un processus de Itô :

$$\begin{cases} dX_{1,t} &= \sigma_{1,t}dB_{1,t} + b_1(t, \omega)dt, \\ dX_{2,t} &= \sigma_{2,t}dB_{2,t} + b_2(t, \omega)dt, \end{cases}$$

où $(b_\ell(t, \omega))_{t \geq 0}$ est un processus adapté représentant la vitesse moyenne (aussi appelé dérive) et B_1 et B_2 sont des mouvements browniens standards corrélés $\text{Corr}(B_{1,t}, B_{2,t}) = \rho_t$. On peut aussi écrire

$dB_{2,t} = \rho_t dB_{1,t} + \sqrt{1 - \rho_t^2} dB_{3,t}$, où $B_1 = (B_{1,t})_{t \in [0,1]}$ et $B_3 = (B_{3,t})_{t \in [0,1]}$ sont des mouvements browniens indépendants.

On limite notre attention au cas où σ_1 , σ_2 et ρ sont des fonctions déterministes. Les fonctions σ_ℓ , $\ell = 1, 2$ prennent des valeurs positives tandis que ρ prend ses valeurs dans l'intervalle $] -1, 1[$.

Le problème abordé ici est le suivant : nous supposons que $\sigma_{\ell,t}$ et ρ_t sont inconnues et nous cherchons à les estimer à partir d'un échantillon de $(X_t)_{0 \leq t \leq 1}$, ou, plus exactement, nous estimons le vecteur de la (co)-variation inconnu de la diffusion X , défini par

$$V_t = ([X_1]_t, [X_2]_t, \langle X_1, X_2 \rangle_t)^T,$$

quand $X_{1,t}$ et $X_{2,t}$ sont observés d'une manière synchrone, $[X_\ell]_t := \int_0^t \sigma_s^2 ds$, $\ell = 1, 2$ représente la variation quadratique du processus X_ℓ et $\langle X_1, X_2 \rangle_t := \int_0^t \sigma_{1,s} \sigma_{2,s} \rho_s ds$ la covariance déterministe de X_1 et X_2 .

Cette question apparaît naturellement en mathématiques financières, V_t étant alors appelé (co)-volatilité. Étant donnés des observations discrètes également espacées $(X_{1,t_k^n}, X_{2,t_k^n}, k = 1, \dots, n)$ dans l'intervalle $[0, 1]$ (avec $t_k^n = k/n$), un estimateur naturel pour V_t est le vecteur de la (co)-variance réalisée

$$V_t^n(X) = (Q_{1,t}^n(X), Q_{2,t}^n(X), C_t^n(X))^T,$$

dans lequel, les quantités $Q_{\ell,t}^n(X)$ et $C_t^n(X)$ pour $\ell = 1, 2$, sont donnés par

$$Q_{\ell,t}^n(X) := \sum_{k=1}^{[nt]} (\Delta_k X_\ell)^2 \quad \text{et} \quad C_t^n := \sum_{k=1}^{[nt]} \Delta_k X_1 \Delta_k X_2,$$

où $\Delta_k X_\ell = X_{\ell,t_k^n} - X_{\ell,t_{k-1}^n}$ avec $\tau_n = \{t_k^n := \frac{k}{n}; 0 \leq k \leq n\}$, une partition régulière de $[0, 1]$ à n morceaux, la notation $[x]$ sera utilisée pour la partie entière de $x \in \mathbb{R}$.

Les théorèmes limite pour les processus stochastiques affirment que $V_t^n(X)$ est un estimateur consistant de V_t . Le TCL est aussi prouvé (voir [150]). L'objectif principal de cette partie est d'obtenir des informations plus fines sur cet estimateur, notamment de raffiner le TCL en nous intéressant aux grandes déviations.

En collaboration avec A. GUILLIN et L. WU [T-D1], nous avons d'abord obtenu le principe de grandes déviations et des déviations modérées pour l'estimateur de la variation quadratique $Q_{\ell,t}^n(X)$ par une approche directe. Nous avons obtenu des inégalités de déviations dans ce cas.

En collaboration avec Y. Samoura [H*-D12], nous avons considéré les grandes déviations pour l'estimateur de la covariance $C_t^n(X)$. La preuve découle d'une application du principe de contraction aux résultats donnés dans le cas de la volatilité.

Récemment avec A. Guillin et Y. Samoura [H-D13], nous avons considéré les grandes déviations pour le vecteur de la variance et de la covariance du processus de diffusion bidimensionnelle $V_t^n(X)$. Nous ne présentons ici que ces derniers résultats.

On introduit la fonction suivante, qui va jouer un rôle crucial dans le calcul de la fonction génératrice des moments : pour $-1 < c < 1$ pour tout $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$

$$(1.1) \quad P_c(\lambda) := \begin{cases} -\frac{1}{2} \log \left(\frac{(1 - 2\lambda_1(1 - c^2))(1 - 2\lambda_2(1 - c^2)) - (\lambda_3(1 - c^2) + c)^2}{1 - c^2} \right) & \text{si } \lambda \in \mathcal{D}_c \\ +\infty, & \text{sinon} \end{cases}$$

où

$$\mathcal{D}_c = \left\{ \lambda \in \mathbb{R}^3, \max_{\ell=1,2} \lambda_\ell < \frac{1}{2(1 - c^2)} \text{ et } \prod_{\ell=1}^2 (1 - 2\lambda_\ell(1 - c^2)) > (\lambda_3(1 - c^2) + c)^2 \right\}.$$

La fonction P_c est la log-Laplace de la variable aléatoire $\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 XY$, où X, Y sont deux variables aléatoires gaussiennes standards corrélées avec un coefficient c .

Introduisons aussi la fonction P_c^* qui est la transformée de Legendre de P_c donnée dans (1.1) : pour tout $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, soit

$$(1.2) \quad P_c^*(x) := \begin{cases} \log \left(\frac{\sqrt{1-c^2}}{\sqrt{x_1 x_2 - x_3^2}} \right) - 1 + \frac{x_1 + x_2 - 2c x_3}{2(1-c^2)} & \text{si } x_1 > 0, x_2 > 0, x_1 x_2 > x_3^2 \\ +\infty, & \text{sinon.} \end{cases}$$

Nous nous intéressons aux grandes déviations de niveau processus (toute la trajectoire est considérée) de $V^n(X)$, cas particulièrement intéressant du point de vue de la statistique non paramétrique.

Soit $\mathcal{BV}([0, 1], \mathbb{R}^3)$ l'espace des fonctions à variations bornées sur $[0, 1]$. On identifie \mathcal{BV} avec $\mathcal{M}_3([0, 1])$, l'espace des mesures vectorielles à valeurs dans \mathbb{R}^3 . Cette identification sera faite de la manière usuelle : pour $f \in \mathcal{BV}$, il lui correspond μ^f caractérisée par $\mu^f([0, t]) = f(t)$. L'espace $\mathcal{C}_3([0, 1])$ des fonctions continues bornées de $[0, 1]$ à valeurs dans \mathbb{R}^3 est l'espace topologique dual de \mathcal{BV} . On munit \mathcal{BV} de la topologie de la convergence faible* $\sigma(\mathcal{BV}, \mathcal{C}_3([0, 1]))$.

Soit $f \in \mathcal{BV}$ et μ^f la mesure associée dans $\mathcal{M}_3([0, 1])$. Considérons la décomposition de Lebesgue de μ^f , $\mu^f = \mu_a^f + \mu_s^f$ où μ_a^f désigne la partie absolument continue de μ^f par rapport à dx et μ_s^f sa partie singulière. On note $f_a(t) = \mu_a^f([0, t])$ et $f_s(t) = \mu_s^f([0, t])$.

Théorème 1.2.1 *Supposons que $b(\cdot, \cdot) \in L^\infty(dt \otimes \mathbb{P})$ que $\sigma_{\ell,t}^2(1 - \rho_t^2)$ et $\sigma_{1,t}\sigma_{2,t}(1 - \rho_t^2) \in L^\infty([0, 1], dt)$ et que les fonctions $t \rightarrow \sigma_{\ell,t}$ et $t \rightarrow \rho_t$ sont continues. Alors*

1. $\mathbb{P}(V_1^n(X) \in \cdot)$ satisfait le PGD sur \mathbb{R}^3 de vitesse n et de bonne fonction de taux donnée par

$$(1.3) \quad I_{ldp}(x) = \sup_{\lambda \in \mathbb{R}^3} \left(\langle \lambda, x \rangle - \int_0^1 P_{\rho_t}(\lambda_1 \sigma_{1,t}^2, \lambda_2 \sigma_{2,t}^2, \lambda_3 \sigma_{1,t} \sigma_{2,t}) dt \right).$$

où la fonction P_c est donnée dans (1.1).

2. $\mathbb{P}(V^n(X) \in \cdot)$ satisfait le PGD sur \mathcal{BV} de vitesse n et de bonne fonction de taux J_{ldp} donnée par

$$\begin{aligned} J_{ldp}(f) &= \int_0^1 P_{\rho_t}^* \left(\frac{f'_{1,a}(t)}{\sigma_{1,t}^2}, \frac{f'_{2,a}(t)}{\sigma_{2,t}^2}, \frac{f'_{3,a}(t)}{\sigma_{1,t} \sigma_{2,t}} \right) dt \\ &\quad + \int_0^1 \frac{\sigma_{2,t}^2 f'_{1,s}(t) + \sigma_{1,t}^2 f'_{2,s}(t) - 2\rho_t \sigma_{1,t} \sigma_{2,t} f'_{3,s}(t)}{2\sigma_{1,t}^2 \sigma_{2,t}^2 (1 - \rho_t^2)} 1_{[t; f'_{1,s} > 0, f'_{2,s} > 0, (f'_{3,s})^2 < f'_{1,s} f'_{2,s}]} d\theta(t), \end{aligned}$$

pour tout $f = (f_1, f_2, f_3) \in \mathcal{BV}$ où P_c^* est donnée dans (1.2) et θ est n'importe quelle mesure à valeurs réelles positives par rapport à laquelle μ_s^f est absolument continue et $f'_s = d\mu_s^f/d\theta = (f'_{1,s}, f'_{2,s}, f'_{3,s})$.

3. On suppose que pour $\ell = 1, 2$, σ_ℓ et ρ sont constants et que $b(\cdot, \cdot) \in L^\infty(dt \otimes \mathbb{P})$. Alors $\mathbb{P}(V_1^n(X) \in \cdot)$ satisfait le PGD sur \mathbb{R}^3 de vitesse n et de bonne fonction de taux I_{ldp}^V donnée par

$$I_{ldp}^V(x_1, x_2, x_3) = P_\rho^* \left(\frac{x_1}{\sigma_1^2}, \frac{x_2}{\sigma_2^2}, \frac{x_3}{\sigma_1 \sigma_2} \right).$$

Notons que la définition de f'_s est θ -dépendante. Cependant, par homogénéité, J_{ldp} ne dépend pas de θ . On peut choisir $\theta = |f_{1,s}| + |f_{2,s}| + |f_{3,s}|$, avec $|f_{i,s}| = f_{i,s}^+ + f_{i,s}^-$, où $f_{i,s} = f_{i,s}^+ - f_{i,s}^-$ est la décomposition de Hahn-Jordan.

Puisque nous ne pouvons pas déterminer le comportement limite de la fonction génératrice des moments à un certain point de la frontière du domaine de définition, nous utilisons une autre approche basée sur les résultats de Najim [134, 135, 133], où l'hypothèse d'escarpement de la fonction génératrice des moments est relaxée.

En utilisant le principe de contraction, nous retrouvons les résultats du cas unidimensionnel.

Nous rappelons que nous avons obtenu le principe de grandes déviations pour $\mathbb{P}(Q^n(X) \in \cdot)$ avec la fonction de taux

$$J(\gamma) = \int_0^1 P^* \left(\frac{\dot{\gamma}(t)}{\sigma_t^2} \right) dt + \frac{1}{2} \int_0^1 \frac{1}{\sigma_t^2} d\gamma^\perp(t) \quad \text{où} \quad P^*(x) = \begin{cases} \frac{1}{2}(x - 1 - \log x) & \text{si } x > 0 \\ +\infty & \text{si } x \leq 0. \end{cases}$$

Notre approche dans ce cas étant directe, notre preuve découle de l'approximation fini-dimensionnelle (soit en temps fixe) obtenue par approximation, et de l'indépendance des accroissements de $Q^n(X)$. L'identification de la fonction de taux donnée comme la limite du taux du PGD fini-dimensionnel, est beaucoup plus délicate. L'élément clé est une réécriture de la fonction $J(\gamma)$ à obtenir. Dans le cas où $\sigma_t = \sigma$ est constante, ce résultat a déjà été démontré par Bercu, Gamboa et Rouault [23] par l'étude d'opérateurs de Toeplitz. Les grandes déviations liées à la topologie faible ne permettent pas d'obtenir des estimations pour des domaines de déviations du type $\{\gamma : \sup_{t \in [0,1]} |\gamma(t) - [X]_t| \geq r\}$, intéressants en statistique. Pour pallier ce problème, les inégalités de déviations suivantes sont démontrées.

Proposition 1.2.2 *Soit $Y := \int_0^\cdot b(t, \omega) dt$. Alors, pour tout $n \geq 1$ et tout $r > 0$,*

$$\mathbb{P} \left(\sup_{t \in [0,1]} [Q_t^n(X - Y) - \mathbb{E} Q_t^n(X - Y)] \geq r \right) \leq e^{-\frac{n}{2} \left[\frac{r}{\|\sigma^2\|_\infty} - \log \left(1 + \frac{r}{\|\sigma^2\|_\infty} \right) \right]}$$

$$\mathbb{P} \left(\inf_{t \in [0,1]} [Q_t^n(X - Y) - \mathbb{E} Q_t^n(X - Y)] \leq -r \right) \leq e^{-n \frac{r^2}{4 \int_0^1 \sigma_t^4 dt}}.$$

Nous rappelons aussi la fonction de taux obtenue pour les grandes déviations de $\mathbb{P}(C^n \in \cdot)$

$$J(\gamma) = J_+^{abs}(\gamma_+ + \beta) + J_-^{abs}(\gamma_- + \beta) + J_+^\perp(\gamma_+) + J_-^\perp(\gamma_-),$$

où β est absolument continue par rapport à la mesure de Lebesgue et est donnée par

$$\begin{aligned} \dot{\beta}(t) &= \frac{\sigma_{1,t}\sigma_{2,t}(1 - \rho_t^2) - (\dot{\gamma}_+(t) + \dot{\gamma}_-(t))}{2} \\ &+ \frac{\sqrt{[\sigma_{1,t}\sigma_{2,t}(1 - \rho_t^2) - (\dot{\gamma}_+(t) + \dot{\gamma}_-(t))]^2 + (\dot{\gamma}_+(t) + \dot{\gamma}_-(t))\sigma_{1,t}\sigma_{2,t}(1 - \rho_t^2)}}{2}, \end{aligned}$$

et

$$J_\pm^\perp(\gamma) = \int_0^1 \frac{1}{\sigma_{1,t}\sigma_{2,t}(1 \pm \rho_t)} d\gamma^\perp, \quad \text{et} \quad J_\pm^{abs}(\gamma) = \int_0^1 P^* \left(\frac{2\dot{\gamma}(t)}{\sigma_{1,t}\sigma_{2,t}(1 \pm \rho_t)} \right) dt.$$

L'hypothèse de continuité de $\sigma_{\ell,\cdot}$ et ρ_\cdot n'est pas nécessaire. En son absence, on doit considérer une autre stratégie pour la preuve, plus technique et basée sur le théorème de Dawson-Gärtner, mais qui ne permet pas d'obtenir une autre précision sur la fonction de taux outre le fait que ce soit une bonne fonction de taux.

Cependant, il n'est pas difficile d'adapter notre preuve au cas où $\sigma_{\ell,\cdot}$ et ρ_\cdot ont seulement un nombre fini de points de discontinuité (du premier type). Ceci peut être fait en appliquant le théorème précédent à chaque sous-intervalle où toutes les fonctions sont continues et en utilisant l'indépendance des incréments.

Considérons maintenant l'échelle intermédiaire entre le TCL et loi des grands nombres. Soit \mathcal{H} l'espace de Banach des fonctions croissantes à valeurs dans \mathbb{R}^3 càdlàg γ sur $[0, 1]$ avec $\gamma(0) = 0$, muni de la norme de la convergence uniforme et de la σ -algèbre \mathcal{B}^s engendrée par les coordonnées $\{\gamma(t), 0 \leq t \leq 1\}$.

Théorème 1.2.3 Supposons $b(\cdot, \cdot) \in L^\infty(dt \otimes \mathbb{P})$ et $\sigma_{\ell,t}^2(1 - \rho_t^2) \in L^2([0, 1], dt)$. Soit $(v_n)_{n \geq 1}$ une suite de nombres réels positifs telle que

$$(1.4) \quad v_n \xrightarrow{n \rightarrow \infty} \infty, \quad \frac{v_n}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} 0, \quad \text{et} \quad \sqrt{n} v_n \max_{\ell=1,2} \max_{1 \leq k \leq n} \int_{(k-1)/n}^{k/n} \sigma_{\ell,t}^2 dt \xrightarrow{n \rightarrow \infty} 0.$$

Cette condition est vérifiée si pour un certain $p > 2$, $\sigma_t^2 \in L^p([0, 1], dt)$ et $b_n = O\left(n^{\frac{1}{2} - \frac{1}{p}}\right)$.

Alors on a

1. $\mathbb{P}(\sqrt{n}(V_1^n(X) - [V]_1)/v_n \in \cdot)$ satisfait le PGD sur \mathbb{R}^3 de vitesse v_n^2 et de fonction de taux

$$I_{mdp}(x) = \sup_{\lambda \in \mathbb{R}^3} \left(\langle \lambda, x \rangle - \frac{1}{2} \langle \lambda, \Sigma_1 \cdot \lambda \rangle \right) = \frac{1}{2} \langle x, \Sigma_1^{-1} \cdot x \rangle, \quad \text{avec} \quad \Sigma_1 = \int_0^1 \bar{\Sigma}_t dt \quad \text{où}$$

$$\bar{\Sigma}_t = \begin{pmatrix} \sigma_{1,t}^4 & \sigma_{1,t}^2 \sigma_{2,t}^2 \rho_t^2 & \sigma_{1,t}^3 \sigma_{2,t} \rho_t \\ \sigma_{1,t}^2 \sigma_{2,t}^2 \rho_t^2 & \sigma_{2,t}^4 & \sigma_{1,t} \sigma_{2,t}^3 \rho_t \\ \sigma_{1,t}^3 \sigma_{2,t} \rho_t & \sigma_{1,t} \sigma_{2,t}^3 \rho_t & \frac{1}{2} \sigma_{1,t}^2 \sigma_{2,t}^2 (1 + \rho_t^2) \end{pmatrix}$$

2. $\mathbb{P}(\sqrt{n}(V^n(X) - [V]) / v_n \in \cdot)$ satisfait le PGD sur \mathcal{H} de vitesse v_n^2 et de fonction de taux

$$(1.5) \quad J_{mdp}(\phi) = \begin{cases} \int_0^1 \frac{1}{2} \left\langle \dot{\phi}(t), \bar{\Sigma}_t^{-1} \cdot \dot{\phi}(t) \right\rangle dt & \text{si } \phi \in \mathcal{AC}_0([0, 1]) \\ +\infty, & \text{sinon,} \end{cases}$$

où $\bar{\Sigma}_t$ donnée précédemment, est inversible et son inverse $\bar{\Sigma}_t^{-1}$ est donné par

$$\bar{\Sigma}_t^{-1} = \frac{1}{\det(\bar{\Sigma}_t)} \begin{pmatrix} \frac{1}{2} \sigma_{1,t}^2 \sigma_{2,t}^6 (1 - \rho_t^2) & \frac{1}{2} \sigma_{1,t}^4 \sigma_{2,t}^4 \rho_t^2 (1 - \rho_t^2) & -\sigma_{1,t}^3 \sigma_{2,t}^5 \rho_t (1 - \rho_t^2) \\ \frac{1}{2} \sigma_{1,t}^4 \sigma_{2,t}^4 \rho_t^2 (1 - \rho_t^2) & \frac{1}{2} \sigma_{1,t}^6 \sigma_{2,t}^2 (1 - \rho_t^2) & -\sigma_{1,t}^5 \sigma_{2,t}^3 \rho_t (1 - \rho_t^2) \\ -\sigma_{1,t}^3 \sigma_{2,t}^5 \rho_t (1 - \rho_t^2) & -\sigma_{1,t}^5 \sigma_{2,t}^3 \rho_t (1 - \rho_t^2) & \sigma_{1,t}^4 \sigma_{2,t}^4 (1 - \rho_t^4) \end{pmatrix},$$

où $\det(\bar{\Sigma}_t) = \frac{1}{2} \sigma_{1,t}^6 \sigma_{2,t}^6 (1 - \rho_t^2)^3$ et $\mathcal{AC}_0 = \{\phi : [0, 1] \rightarrow \mathbb{R}^3 \text{ absolument continue avec } \phi(0) = 0\}$.

Dans les résultats précédents, nous avons imposé la bornitude de $b(t, \omega)$ ce qui permet de réduire facilement les grandes déviations de $V^n(X)$ au cas sans drift. Nos résultats restent vrais sous la condition de Lipschitz ou plus généralement sous une condition de croissance linéaire du drift $b(t, x)$.

Tous les résultats restent valables quand la taille de discrétisation n'est pas régulière. La propriété importante est que chaque point soit dans le support de la mesure limite.

Lorsque le processus de volatilité est aléatoire, mais indépendant du mouvement brownien, nous pouvons obtenir les grandes déviations pour les statistiques du vecteur de la (co)-volatilité réalisée conditionnellement aux trajectoires du processus de la volatilité. Les preuves peuvent facilement être modifiées pour obtenir les principaux résultats par l'interprétation de la probabilité et de l'espérance mathématique comme étant conditionnelles à la donnée de la trajectoire du processus de la volatilité, voir Kanaya et Otsu [109].

Applications : PGD pour les coefficients réalisés de corrélation et de régression

Dans cette partie, nous appliquons nos résultats pour obtenir le PGD et le PDM pour des mesures standards de dépendances entre deux retours d'actifs comme le coefficient de régression réalisée $\beta_{\ell,1} = C_1/Q_{\ell,1}$ pour $\ell = 1, 2$ et la corrélation réalisée $\varrho_1 = C_1/\sqrt{Q_{1,1}Q_{2,1}}$ qui sont estimés par $\beta_{\ell,1}^n(X) = C_1^n(X)/Q_{\ell,1}^n(X)$ et $\varrho_1^n(X) = C_1^n(X)/\sqrt{Q_{1,1}^n(X)Q_{2,1}^n(X)}$ respectivement. Soulignons que les grandes déviations pour la corrélation réalisée ne peuvent pas être déduites à partir du cas unidimensionnel et ont donc été largement ignorées. L'application est basée essentiellement sur la delta méthode, développée par Gao et Zhao ([85]). Pour simplifier, nous nous concentrons sur le cas où σ_ℓ pour $\ell = 1, 2$ sont constants et nous notons par $\varrho := \int_0^1 \rho_t dt$.

Parmi les travaux récents qui ont considéré ces estimateurs, on peut citer Andersen, Bollerslev, Diebold et Wu [5], Todorov et Bollerslev [164], Dovonon, Gonçalves et Meddahi [69]. La consistance et le TCL pour ces estimateurs ont déjà été étudiés, voir par exemple Mancini et Gobbi [123].

Proposition 1.2.4 *Soit σ_ℓ constante et $\varrho := \int_0^1 \rho_t dt$. Sous les conditions du Théorème 1.2.1,*

1. $\mathbb{P}(\varrho_1^n(X) \in \cdot)$ *satisfait le PGD sur \mathbb{R} de vitesse n et de bonne fonction de taux donnée par*

$$I_{ldp}^\varrho(u) = \inf_{\{(x,y,z) \in \mathbb{R}^3 : u = \frac{z}{\sqrt{xy}}\}} I_{ldp}(x, y, z)$$

où I_{ldp} est donnée dans (1.3). Si ρ est constant, on obtient

$$(1.6) \quad I_{ldp}^\rho(u) = \begin{cases} \log \left(\frac{1 - \rho u}{\sqrt{1 - \rho^2} \sqrt{1 - u^2}} \right), & \text{si } -1 < u < 1 \\ +\infty, & \text{sinon.} \end{cases}$$

2. *pour $\ell = 1$ ou 2 , $\mathbb{P}(\beta_{\ell,1}^n(X) \in \cdot)$ satisfait le PGD sur \mathbb{R} de vitesse n et de fonction de taux*

$$I_{ldp}^{\beta_{\ell,1}}(u) = \inf_{\{(x_1, x_2, x_3) \in \mathbb{R}^3 : u = \frac{x_3}{x_\ell}\}} I_{ldp}(x_1, x_2, x_3)$$

où I_{ldp} est donnée par (1.3). Si ρ est constant, on obtient

$$(1.7) \quad I_{ldp}^{\beta_\ell}(u) = \frac{1}{2} \log \left(1 + \frac{(\sigma_\ell u - \rho \sigma_\iota)^2}{\sigma_\iota^2 (1 - \rho^2)} \right).$$

Comme on peut l'imaginer en regardant la forme de l'expression de la fonction de taux, ceci n'est qu'une simple application du principe de contraction à partir du PGD du vecteur de la volatilité.

La preuve du PDM est beaucoup plus difficile et nécessite une technologie un peu plus subtile : grandes déviations via la delta méthode développée par Gao et Zhao ([85]).

Proposition 1.2.5 *Soit σ_ℓ constant et $\varrho := \int_0^1 \rho_t dt$. Sous les conditions du Théorème 1.2.3,*

1. $\mathbb{P}(\sqrt{n}(\varrho_1^n(X) - \varrho)/v_n \in \cdot)$ *satisfait le PGD sur \mathbb{R} de vitesse v_n^2 et de fonction de taux*

$$I_{mdp}^\varrho(u) = \inf_{\{(x,y,z) \in \mathbb{R}^3 : u = \frac{z}{\sigma_1 \sigma_2} - \varrho \frac{\sigma_1^2 y + x \sigma_2^2}{2 \sigma_1^2 \sigma_2^2}\}} I_{mdp}(x, y, z).$$

Si ρ est constant, on obtient $I_{mdp}^\rho(u) = \frac{u^2}{(1 - \rho^2)^2}$.

2. *pour $\ell, \iota \in \{1, 2\}$ avec $\ell \neq \iota$, $\mathbb{P}(\sqrt{n}(\beta_{\ell,1}^n(X) - \varrho \frac{\sigma_\iota}{\sigma_\ell})/v_n \in \cdot)$ satisfait un PGD sur \mathbb{R} de vitesse v_n^2 et de fonction de taux donnée par*

$$I_{mdp}^{\beta_{\ell,1}}(u) = \inf_{\{(x,y,z) \in \mathbb{R}^3 : u = \frac{z}{\sigma_\ell^2} - \varrho \frac{\sigma_\iota}{\sigma_\ell^2} x\}} I_{mdp}(x, y, z).$$

Si ρ est constant, on obtient $I_{mdp}^{\beta_{\ell,1}}(u) = \frac{\sigma_\ell^2 u^2}{\sigma_\iota^2 (1 - \rho^2)}$.

1.3 Grandes déviations dans le cas synchrone avec sauts

Les processus avec sauts sont des outils standards, par exemple, pour modéliser des valeurs d'un actif financier ou dans l'assurance. La motivation clé derrière ces modèles de diffusion avec sauts est la modélisation du marché des actions présentant des changements soudains du prix, ou celle d'un titre risqué qui ne peut pas être modélisé par une seule composante diffusive. L'utilisation de sauts dans les modèles financiers semble donc être de plus en plus nécessaire pour les applications pratiques.

Sur un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{[0,1]}, \mathbb{P})$, on considère $X_1 = (X_{1,t})_{t \in [0,1]}$ et $X_2 = (X_{2,t})_{t \in [0,1]}$ deux processus définis par une diffusion avec sauts de Lévy construit *via* une superposition d'un processus de Wiener avec drift, et d'une composante de Processus de Poisson indépendante.

Plus précisément, $X_1 = (X_{1,t})_{t \in [0,1]}$ et $X_2 = (X_{2,t})_{t \in [0,1]}$ sont donnés par

$$\begin{cases} dX_{1,t} = b_1(t, \omega)dt + \sigma_{1,t}dB_{1,t} + dJ_{1,t} \\ dX_{2,t} = b_2(t, \omega)dt + \sigma_{2,t}dB_{2,t} + dJ_{2,t} \end{cases}$$

pour $t \in [0,1]$ où $B_{1,t}$ et $B_{2,t}$ sont deux mouvement browniens corrélés, avec $d\langle B_1, B_2 \rangle_t = \rho_t dt$. Les processus $J_{1,t}$ et $J_{2,t}$ sont des processus de sauts éventuellement corrélés. On suppose que $J_{1,t}$ $J_{2,t}$ sont à activité finie, c.à.d. qu'il y a un nombre fini de sauts sur un intervalle fini. Un modèle de Lévy général doit contenir une composante compensée de sauts infinie.

Sous nos hypothèses, J_ℓ est nécessairement un processus de Poisson composé et donc peut-être réécrit comme $J_{\ell,t} = \sum_{i=1}^{N_{\ell,t}} Y_{\ell,i}$. Ici les variables aléatoires $Y_{\ell,i}$ sont i.i.d. de loi ν_ℓ/λ_ℓ , où ν_ℓ est la mesure de Lévy de X_ℓ normalisé par la masse totale $\lambda_\ell = \nu_\ell(\mathbb{R} - \{0\}) < \infty$. On suppose que N_ℓ est un processus de Poisson indépendant de chaque $Y_{\ell,i}$ d'intensité constante λ_ℓ .

Au cours des dernières années, plusieurs méthodes d'estimation du vecteur de la (co-)volatilité ont été proposées. Nous utilisons l'estimateur tronqué par une fonction de seuil qui est introduit indépendamment par Mancini [121] et par Shimizu et Yoshida [159]. Dans cette méthode, seules les variations inférieures à une fonction de seuil donnée sont prises en compte.

Ainsi, étant donné l'observation synchrone et également espacée du processus $X_{1,t_0}, X_{1,t_1}, \dots, X_{1,t_n}, X_{2,t_0}, X_{2,t_1}, \dots, X_{2,t_n}$, nous considérons la statistique définie précédemment. Cependant cet estimateur est fortement influencé quand les processus X_ℓ contiennent des sauts. En fait, quand n tend vers l'infini, une telle somme s'approche du vecteur variance-covariation quadratique globale $([X_1]_t, [X_2]_t, [X_1, X_2]_t)$ qui contient aussi les co-sauts, où

$$[X_\ell]_t := \int_0^t \sigma_{\ell,s}^2 ds + \sum_{s \leq t} (\Delta J_{\ell,s})^2, \quad [X_1, X_2]_t := \int_0^t \sigma_{1,s} \sigma_{2,s} \rho_s ds + \sum_{s \leq t} \Delta J_{1,s} \Delta J_{2,s} \quad \text{et} \quad \Delta J_{\ell,s} = J_{\ell,s} - J_{\ell,s^-}.$$

Mais si nous prenons une fonction déterministe $r(\frac{1}{n})$, au pas $\frac{1}{n}$ entre les observations, telle que

$$\lim_{n \rightarrow \infty} r\left(\frac{1}{n}\right) = 0, \quad \text{et} \quad \lim_{n \rightarrow \infty} \frac{\log n}{nr\left(\frac{1}{n}\right)} = 0,$$

nous pouvons récupérer $V_t = \left(\int_0^t \sigma_{1,s}^2 ds, \int_0^t \sigma_{2,s}^2 ds, \int_0^t \sigma_{1,s} \sigma_{2,s} \rho_s ds \right)$ par l'utilisation de l'estimateur tronqué suivant $\mathcal{V}_t^n = (\mathcal{Q}_{1,t}^n, \mathcal{Q}_{2,t}^n, \mathcal{C}_t^n)$ où

$$\mathcal{Q}_{\ell,t}^n = \sum_{k=1}^{[nt]} (\Delta_k X_\ell)^2 \mathbf{1}_{\{(\Delta_k X_\ell)^2 \leq r(\frac{1}{n})\}} \quad \text{et} \quad \mathcal{C}_t^n = \sum_{k=1}^{[nt]} \Delta_k X_1 \Delta_k X_2 \mathbf{1}_{\{\max_{\ell=1}^2 (\Delta_k X_\ell)^2 \leq r(\frac{1}{n})\}}.$$

La fonction $r(\cdot)$ est une fonction de seuil, telle que quand $|\Delta_k X_\ell| > r(\frac{1}{n})$ un saut a dû se produire dans $]t_{k-1}, t_k]$.

Depuis les travaux fondateurs de Mancini [121], plusieurs auteurs ont exploité ou étendu le concept de seuil pour faire face à des modèles stochastiques complexes, voir Shimizu et Yoshida [159], ou Ogihara

et Yoshida [140]. La même idée de troncature est également utilisée par divers auteurs dans différents contextes, voir par exemple Aït-Sahalia et al. [1], [3] et [4], Gobbi et Mancini [88], Cont et Mancini [123]. Dans [78], les auteurs déterminent ce qui constitue une bonne séquence de seuil r_n et ils proposent une méthode objective pour sélectionner une telle séquence.

Dans ce cas, Mancini [123] a montré que \mathcal{V}_t^n est un des estimateurs convergents de V_t et présente une certaine normalité asymptotique. En outre, quand $\sigma_t = \sigma$, Mancini [122] a étudié les grandes déviations pour l'estimateur tronqué. Jiang [107] a obtenu le principe des déviations modérées fonctionnel pour l'estimateur tronqué. Dans cette partie, nous considérons les grandes déviations pour l'estimateur tronqué du vecteur.

Théorème 1.3.1

1. Supposons $b(\cdot, \cdot) \in L^\infty(dt \otimes \mathbb{P})$ et $\sigma_{\ell,t}^2(1 - \rho_t^2) \in L^\infty([0, 1], dt)$ et les fonctions $t \rightarrow \sigma_{\ell,t}$ et $t \rightarrow \rho_t$ continues.

Soit r telle que $r \left(\frac{1}{n} \right) \xrightarrow{n \rightarrow \infty} 0$ et $nr \left(\frac{1}{n} \right) \xrightarrow{n \rightarrow \infty} \infty$.

Alors on a le PGD du Théorème 1.2.1.

2. Supposons $b(\cdot, \cdot) \in L^\infty(dt \otimes \mathbb{P})$ et $\sigma_{\ell,t}^2(1 - \rho_t^2) \in L^2([0, 1], dt)$.

Soit $(v_n)_{n \geq 1}$ des nombres positifs tels que $v_n \xrightarrow{n \rightarrow \infty} \infty$ et $\frac{v_n}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} 0$ et $\sqrt{n}v_n r \left(\frac{1}{n} \right) = O(1)$ et

$$r \left(\frac{1}{n} \right) / \left(\log \left(\frac{n}{v_n^2} \right) \max_{k=1}^n \int_{t_{k-1}}^{t_k} \sigma_{\ell,s}^2 ds \right) \longrightarrow +\infty.$$

Alors on a le PDM du Théorème 1.2.3.

La preuve est basée sur les techniques d'approximation au sens des grandes déviations.

Sous la condition $b_\ell = 0$, on peut prouver que pour tout $\theta \in \mathbb{R}^3$:

$$\lim_{n \rightarrow \infty} \frac{1}{v_n^2} \log \mathbb{E} \left(e^{\sqrt{n}v_n \langle \theta, \mathcal{V}_1^n - V_1 \rangle} \right) = \frac{1}{2} \langle \theta, \Sigma_1 \cdot \theta \rangle.$$

Ceci donne une méthode alternative à la preuve des déviations modérées, classiquement obtenues en utilisant le théorème de Gärtner-Ellis.

On peut aussi calculer la fonction génératrice des moments de $\mathcal{V}_1^n(X)$. On obtient pour tout $\theta = (\theta_1, \theta_2, \theta_3)^T \in \mathcal{D}_{\rho_t}$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left(e^{n \langle \theta, \mathcal{V}_1^n \rangle} \right) = \int_0^1 P_{\rho_s} (\theta_1 \sigma_{1,s}^2, \theta_2 \sigma_{2,s}^2, \theta_3 \sigma_{1,s} \sigma_{2,s}) ds.$$

Mais l'étude de la condition d'escarpement est plus difficile.

1.4 Grandes déviations dans le cas non-synchrone sans saut

Dans l'estimation de la volatilité en finance, on dispose souvent d'observations à des temps non régulièrement espacés. En effet, dans les applications financières, les données de transactions réelles sont enregistrées à des heures irrégulières d'une manière non synchrone, c.à.d. que deux prix d'une transaction ne sont généralement pas observés aux mêmes instants.

Supposons avoir des observations non-synchrones $(X_{1,s_i}, X_{2,t_j})_{i=0,\dots,n; j=0,\dots,m}$ avec $0 = s_0 < s_1 < \dots < s_n = T, 0 = t_0 < t_1 < \dots < t_m = T, m, n \in \mathbb{N}$. Récemment, Hayashi et Yoshida [99] ont introduit l'estimateur suivant

$$U_{n,m} := \sum_{i=1}^n \sum_{j=1}^m \Delta X_1(I^i) \Delta X_2(J^j) I_{\{I^i \cap J^j \neq \emptyset\}},$$

où $I^i = (s_{i-1}, s_i]$, $J^j = (t_{j-1}, t_j]$. Sous certaines hypothèses, ils ont démontré la consistance et le TCL pour $U_{n,m}$. Dans ce travail en collaboration avec H. Jiang, A. Guillin et Y. Samoura [H-D15], nous nous

intéressons à l'étude des déviations modérées de $U_{n,m}$ et des grandes déviations dans des cas particuliers.

Tout d'abord, un autre design réduit par rapport à $(I^i)_{i=1,\dots,n}$ sera construit comme suit. Nous recueillons les intervalles I^i tels que $I^i \subset J^j$ et nous les combinons en un nouvel intervalle. Si aucun tel intervalle I^i n'existe, on ne fait rien. Ensuite, la collecte de tous ces intervalles et leur ré-étiquetage de gauche à droite donnent une nouvelle partition de $(0, T]$, notée $(\hat{I}^i)_{i=1,\dots,\hat{n}}$.

En utilisant la bilinéarité de $U_{n,m}$,

$$U_{n,m} = \sum_{i=1}^{\hat{n}} \sum_{j=1}^m \Delta X_1(\hat{I}^i) \Delta X_2(J^j) K_{ij}^{\hat{I}} = \sum_{i=1}^{\hat{n}} \Delta X_1(\hat{I}^i) \Delta X_2 \left(\bigcup_{j \in \hat{J}(i)} J^j \right),$$

où $\hat{J}(i) := \{1 \leq j \leq m : K_{ij}^{\hat{I}} \neq \emptyset\}$ avec $K_{ij}^{\hat{I}} = I_{\{\hat{I}^i \cap J^j \neq \emptyset\}}$.

Ensuite, nous pouvons constater que chaque J^j contient au plus un \hat{I}^i , ce qui implique que la suite de variables aléatoires $\{1 \leq i \leq \hat{n} : \Delta X_1(\hat{I}^i) \Delta X_2 \left(\bigcup_{j \in \hat{J}(i)} J^j \right)\}$ est 2-dépendante.

Pour tout Borélien de $I \subset [0, T]$, on définit $\nu(I) = \int_I \sigma_{1,t} \sigma_{2,t} \rho_t dt$, $\nu_\ell(I) = \int_I \sigma_{\ell,t}^2 dt$, $\ell = 1, 2$. Soit $r_{n,m} := \max_{1 \leq i \leq n} |I^i| \vee \max_{1 \leq j \leq m} |J^j|$, la plus grande taille d'intervalle. On suppose que quand $n, m \rightarrow +\infty$ alors $r_{n,m} \rightarrow 0$.

Théorème 1.4.1 *Sous les hypothèses :*

1. *il existe une suite positive de nombres réels $(b_n)_{n \geq 1}$ et une certaine constante $\Sigma \in (0, +\infty)$ telle que, quand $n \rightarrow +\infty$, $b_n \rightarrow 0$ et*

$$b_n^{-1} \left(\sum_{i=1}^{\hat{n}} \sum_{j=1}^m \nu_1(\hat{I}^i) \nu_2(J^j) K_{ij}^{\hat{I}} + \sum_{i=1}^{\hat{n}} \nu^2(I^i) + \sum_{j=1}^m \nu^2(J^j) - \sum_{i=1}^{\hat{n}} \sum_{j=1}^m \nu^2(I^i \cap J^j) \right) \rightarrow \Sigma.$$

2. *il existe une suite positive de nombres réels positifs $(a_n)_{n \geq 1}$ telle que, quand $n \rightarrow +\infty$,*

$$a_n \rightarrow +\infty, \quad a_n b_n^{1/2} \rightarrow 0 \quad \text{et} \quad a_n b_n^{-1/2} \nu_n \rightarrow 0,$$

$$\text{où } \nu_n = \max_{\ell=1,2} \max_{1 \leq i \leq \hat{n}} \max_{1 \leq j \leq m} (\nu_\ell(I^i) \vee \nu_\ell(J^j));$$

la suite $\left\{ \left(U_{n,m} - \int_0^T \sigma_{1,t} \sigma_{2,t} \rho_t dt \right) / a_n b_n^{1/2}, n \geq 1 \right\}$ satisfait le PGD de vitesse a_n^2 et de fonction de taux $L(x) = x^2 / 2\Sigma$.

La preuve est basée sur l'utilisation de l'approche des déviations modérées des variables aléatoires 1-dépendantes de Chen [46]. Le principe des grandes déviations étant plus difficile dans le cadre général, il est prouvé dans des cas très particuliers. Les résultats de cette partie sont encore très préliminaires.

Problèmes et perspectives

En lien direct avec les travaux précédemment cités sur les grandes déviations des estimateurs du coefficient de diffusion, quelques directions particulières sont l'objet de mes préoccupations. Dans chaque cas, les difficultés à surmonter ne sont pas de même nature et les méthodes sont différentes. Nous espérons que notre travail contribuera à la récente tendance dans la recherche pour des problèmes de l'estimation de la (co-)variance, qui sont assez souvent discutés dans l'analyse de données financières de hautes fréquence. Nous proposons d'étudier les grandes déviations de la (co-)volatilité réalisée, dans des situations très larges (processus à sauts, en présence d'un bruit, asynchrone) et également lorsque l'équation différentielle stochastique ne sera plus dirigée par un processus de Wiener mais par un processus de Lévy ou un mouvement Brownien fractionnaire.

On considère l'EDS suivante

$$dX_t = \sigma_t dB_t + b(t, \omega) dt$$

Nous nous intéressons aux différentes approches pour l'estimation de la (co-)volatilité.

1. L'approche multi-échelle développée par Zhang et ces co-auteurs [183], [184], [185]. On définit

$$(1.8) \quad Y_t = X_t + \varepsilon_t,$$

où ε_t est une variable aléatoire gaussienne (bruit de la microstructure). On répartit l'échantillon observé en plusieurs sous-échantillons. On pose

$$Y_{t_k^n} = X_{t_k^n} + \varepsilon_{t_k^n}, \quad [Y]^{(n,K)} = \frac{1}{K} \sum_{i \geq K} (Y_{t_i^n} - Y_{t_{i-K}^n})^2.$$

Nous avons que

$$\langle X \rangle^{(n)} = \sum_{i=1}^{M_n} \alpha_i [Y]^{(n,K_i)} \rightarrow \int_0^T \sigma_t^2 dt.$$

On choisit α_i et M_n pour obtenir un estimateur nonbiaisé de taux optimal $n^{-1/4}$.

2. L'approche à noyaux non linéaires développée par Barndorff-Nielsen [17], [16]. L'estimateur est basé sur Y donné dans (1.8). Il est défini par

$$Q^n(Y) = \sum_{h=-H}^H \underbrace{K\left(\frac{h}{H+1}\right)}_{\text{Noyau de Parzen}} \gamma_h, \quad \gamma_h = \sum_{j=|h|+1}^n Y_j Y_{j-|h|}, \quad H_n \propto n^{3/5}.$$

3. L'approche de pré-moyennisation, voir Podolskij et Vetter [143], Jacod et ses co-auteurs [105]. L'estimateur est basé sur Y donné dans (1.8). Il est défini par

$$V_t^n = \sum_{i=0}^{[nt]-k_n} |\overline{Y}_i^n|^2, \quad \overline{Y}_i^n = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right) \Delta_{i+j}^n Y, \quad k_n \propto \sqrt{n},$$

où g est une certaine fonction donnée.

4. Le schéma asynchrone initié par Hayashi et Yoshida [99], voir aussi Dalalyan et Yoshida [51].
5. le schéma des observations en temps aléatoires, voir le travail de Fukasawa [81]. On s'intéresse à l'estimation du coefficient de diffusion quand on dispose d'observations à des temps aléatoires. Dans ce cas, la consistance et la normalité asymptotique sont déjà démontrées. Dans ce travail, on étudiera les grandes déviations.

$$Q_T^r(X) := \sum_{j \geq 1, \tau_j \leq T} (X_{\tau_{j+1}} - X_{\tau_j})^2,$$

$\tau = \{\tau_j\}$ suite \nearrow de t.a..

6. Le schéma avec des erreurs d'arrondis, voir les résultats de Delattre et Jacod [59]. On se propose d'étudier le problème des grandes déviations en présence d'erreurs d'arrondis. En effet, en pratique, on ne peut pas connaître exactement les valeurs $X_{t_k^n}$ du fait des erreurs de mesures; on peut considérer ces erreurs comme des erreurs d'arrondis. Plus précisément, que deviennent nos résultats, si on remplace dans les estimateurs de la variation quadratique $\int_0^t \sigma_s^2 ds$, les observations "discrétisées" $X_{t_k^n}$, par des valeurs "arrondies" à un certain niveau de précision $\alpha_n > 0$ (suite de réels strictement positifs). On s'intéresse aux grandes déviations de l'estimateur suivant :

$$Q_t^{(\alpha_n)}(X) := \sum_{k=1}^{[nt]} \left(X_{t_k^n}^{(\alpha_n)} - X_{t_{k-1}^n}^{(\alpha_n)} \right)^2$$

où $X_{t_k^n}^{(\alpha_n)} := \alpha_n \left[\frac{X_{t_k^n}}{\alpha_n} \right]$, $[x]$ désignant la partie entière du réel x . Ce travail est motivé par les résultats de S. Delattre, J. Jacod et M. Rosenbaum.

7. Notre but est de poursuivre les investigations des grandes déviations dans le cas des estimateurs du type

$$n^{\frac{r+s}{2}-1} \sum_{k=2}^{[nt]} |\Delta_k X|^r |\Delta_{k-1} X|^s,$$

qui suscitent beaucoup d'intérêt en finance. La consistance ou la normalité asymptotique de ces estimateurs sont largement étudiées. La méthode pour obtenir les grandes déviations est certainement différente de celle du cas de la variation quadratique, le calcul direct de la transformée de Laplace ne pouvant pas se faire. Le passage par les chaînes de Markov symétriques semble, pour l'instant, l'approche la plus adaptée. On peut généraliser cette approche à des estimateurs du type :

$$n^{-1} \sum_{k=2}^{[nt]} g(\sqrt{n} \Delta_k X) h(\sqrt{n} \Delta_{k-1} X),$$

pour des fonctions g et h assez générales. Diverses approches sont possibles pour les déviations modérées.

Chapitre 2

Chaînes de Markov bifurcantes, théorèmes limites

Ce chapitre présente les résultats obtenus dans [H-D8] publié dans *Annals of Applied Probability* et [H-D10] publié dans *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*.

2.1 Introduction

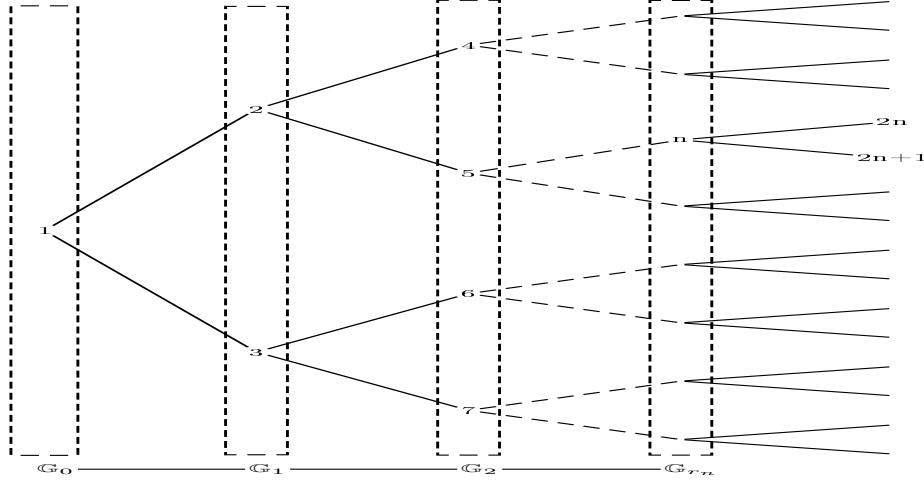
Ce travail a pour origine des expériences effectuées par une équipe de biologistes (Stewart et al. [160]) sur la bactérie *Escherichia coli*. Cette bactérie est un organisme unicellulaire qui se reproduit en se divisant en milieu. Au cours de la division, deux nouvelles membranes ou pôles sont créés, chacun d'eux constituant l'une des extrémités de chacune des filles. Une cellule a donc un pôle "jeune" et un pôle plus "âgé". Lors de la division suivante, on peut donc distinguer les deux nouvelles filles : l'une a hérité du pôle jeune de sa mère et l'autre du pôle plus âgé. Et ainsi de suite. Ce phénomène est schématisé sur la FIGURE 1. Il est possible de suivre et de marquer ainsi toute une généalogie de cellules, de les indexer par un arbre binaire. Nous choisissons l'étiquette $2n$ (resp. $2n + 1$) pour la fille "jeune" (resp. "âgée") de la cellule n . A noter que les cellules dont le label est une puissance de deux (2^n) sont d'une lignée qui a toujours hérité du pôle jeune de sa mère alors que les cellules $2^n + 1$ ont toujours hérité du pôle "âgé", la généalogie des autres cellules est donnée par leur écriture en base deux. Si l'organisme vieillit, alors ce vieillissement se traduit par une dissymétrie entre les deux cellules filles que l'on mesure dans la loi de reproduction elle-même (les cellules paires ont une probabilité plus forte d'avoir deux filles, ...) ou par des caractéristiques physiologiques différentes (taux de croissance, biomasse, longueur, etc.).

Guyon [96] a modélisé l'évolution du taux de croissance par une chaîne de Markov asymétrique sur un arbre binaire régulier, encore appelée chaîne de Markov bifurcante. Ce modèle permet de prendre en compte une répartition asymétrique du taux de croissance de la cellule entre ses deux filles selon leur type. A partir de là, Guyon [96] a construit un test pour détecter une différence entre les taux de croissance des deux filles selon leur type de pôle. Dans ce modèle, on suppose que les cellules ne meurent jamais (la mort signifie ici que la cellule ne se divise plus).

Nous décrivons la généalogie des cellules par un arbre binaire régulier \mathbb{T} dans lequel on voit chaque sommet comme un entier positif différent de 0. La cellule initiale est numérotée 1, elle donne naissance à deux filles, 2 et 3. Plus généralement, la cellule n donne naissance à deux filles numérotées $2n$ et $2n + 1$, et par convention, $2n$ désignera la cellule de type nouveau pôle, $2n + 1$ la cellule de type ancien pôle. Les ensembles

$$\mathbb{G}_n = \{2^n, 2^n + 1, \dots, 2^{n+1} - 1\}, \quad \mathbb{T}_n = \bigcup_{q=0}^n \mathbb{G}_q,$$

désignent alors respectivement la n -ème génération et le sous-arbre des $n + 1$ premières générations. Le taux de croissance de la cellule n est X_n .

FIGURE 2.1 – L'arbre binaire \mathbb{T}

Alors, le cardinal $|\mathbb{G}_n|$ de \mathbb{G}_n est 2^n et celui de \mathbb{T}_n est $|\mathbb{T}_n| = 2^{n+1} - 1$.

Soit \mathbb{G}_{r_n} la génération de l'individu n , ce qui veut dire que $r_n = \lfloor \log_2(n) \rfloor$ où $\lfloor x \rfloor$ désigne la partie entière de x . Rappelons que les deux descendants de l'individu n sont notés $2n$ et $2n+1$, ou inversement, la mère de l'individu n est $\lfloor n/2 \rfloor$. Plus généralement, les ancêtres de l'individu n sont $\lfloor n/2 \rfloor, \lfloor n/2^2 \rfloor, \dots, \lfloor n/2^{r_n} \rfloor$.

Guyon & Al. ([97], [96]) ont proposé le modèle gaussien linéaire suivant pour décrire l'évolution du taux de croissance de la population des cellules tirées d'un individu initial

$$(2.1) \quad \mathcal{L}(X_1) = \nu, \quad \text{et} \quad \forall n \geq 1, \quad \begin{cases} X_{2n} = \alpha_0 X_n + \beta_0 + \varepsilon_{2n} \\ X_{2n+1} = \alpha_1 X_n + \beta_1 + \varepsilon_{2n+1}, \end{cases}$$

où X_n est le taux de croissance de l'individu n , ν est une probabilité sur \mathbb{R} , $\alpha_0, \alpha_1 \in (-1, 1)$, $\beta_0, \beta_1 \in \mathbb{R}$ et $((\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1)$ sont des variables aléatoires i.i.d. de loi $\mathcal{N}_2(0, \Gamma)$, avec

$$(2.2) \quad \Gamma = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \sigma^2 > 0, \quad \rho \in (-1, 1).$$

Le processus (X_n) défini par (2.1) est un exemple typique de BMC qui appelé le processus autorégressif bifurcant du premier ordre (BAR(1)). Les processus BAR(1) sont une adaptation de processus autorégressifs, quand les données ont une arborescence binaire. Ils ont été d'abord présentés par Cowan et Staudte [50], pour des données d'une lignée cellulaire où chaque individu dans une génération donne lieu à deux descendants dans la génération suivante.

Après avoir établi les premiers résultats sur les BMC, Guyon dans [96], a prouvé la loi de logarithme itéré et le TCL pour l'estimateur des moindres carrés $\hat{\theta}^r = (\hat{\alpha}_0^r, \hat{\beta}_0^r, \hat{\alpha}_1^r, \hat{\beta}_1^r)$ du paramètre inconnu $\theta = (\alpha_0, \beta_0, \alpha_1, \beta_1)$. Il a aussi donné quelques tests statistiques permettant de déterminer si le modèle est symétrique ou non (c.à.d. si $\alpha_0 = \alpha_1$ ou non).

Une extension des BMC a été proposée dans [60], dans ce travail les auteurs ont étudié un modèle de BMC avec des données manquantes. Pour prendre en compte la possibilité pour une cellule de mourir, les auteurs dans [60] utilisent l'arbre de Galton-Watson au lieu d'un arbre régulier. Et ils donnent une loi faible des grands nombres, un principe d'invariance et le résultat de TCL pour la moyenne au cours d'une génération ou jusqu'à une génération.

Bercu et al. [22] ont introduit les processus autorégressifs bifurcants d'ordre p . Ils ont fait une analyse assez fine de leur modèle en utilisant les techniques de martingales.

On peut aussi mentionner le travail de De Saporta & al. [56] traitant des processus autorégressifs bifurcants avec des données manquantes dans la procédure d'estimation des paramètres du processus de BAR asymétrique. Ils utilisent un processus de Galton-Watson bi-type pour modéliser la généalogie et pour montrer la convergence et la normalité asymptotique des estimateurs.

Il faut remarquer que l'étude non-asymptotique des inégalités de déviations n'a pas été envisagée du tout dans ces travaux, malgré son intérêt pratique.

Nos objectifs dans ce travail sont multiples. On complète d'abord les travaux de [96], en donnant quelques théorèmes limites pour les BMC (LLN, LIL,...), puis on donne quelques inégalités de déviations pour les moyennes empiriques du type

$$\overline{M}_{\mathbb{T}_r}(f) = \frac{1}{|\mathbb{T}_r|} \sum_{i \in \mathbb{T}_r} f(X_i, X_{2i}, X_{2i+1}),$$

et enfin on étudie les déviations modérées pour les BMC. Comme application de ces résultats, on obtient les déviations modérées et les inégalités de déviations pour l'estimateur des paramètres inconnus dans le BAR(1). Ces résultats sont importants pour une étude statistique rigoureuse. Nous faisons aussi l'extension à des BAR(p) avec p quelconque, en utilisant les outils de martingales.

2.2 Théorèmes limites pour des Chaînes de Markov Bifurcantes

Pour un individu $n \in \mathbb{T}$, on s'intéresse à la quantité X_n (qui peut représenter le poids, le taux de croissance, ...) à valeurs dans un espace métrique S muni de la σ -algèbre \mathcal{S} .

Définition 2.2.1 (\mathbb{T} -probabilité de transition, voir ([96])) *On appelle \mathbb{T} -probabilité de transition toute application $P : S \times \mathcal{S}^2 \rightarrow [0, 1]$ telle que*

- $P(\cdot, A)$ est mesurable pour tout $A \in \mathcal{S}^2$,
- $P(x, \cdot)$ est une mesure de probabilité sur $(\mathcal{S}^2, \mathcal{S}^2)$ pour tout $x \in S$.

Pour une \mathbb{T} -probabilité de transition P sur $S \times \mathcal{S}^2$, on note P_0, P_1 et Q respectivement la première et la seconde marginale de P , et la moyenne de P_0 et P_1 , c.à.d. $P_0(x, B) = P(x, B \times S)$, $P_1(x, B) = P(x, S \times B)$ pour tout $x \in S$ et $B \in \mathcal{S}$ et $Q = \frac{P_0 + P_1}{2}$.

Pour $p \geq 1$, on note $\mathcal{B}(S^p)$ (resp. $\mathcal{B}_b(S^p)$), l'ensemble de toutes les applications S^p -mesurables (resp. S^p -mesurables et bornées) $f : S^p \rightarrow \mathbb{R}$. Pour $f \in \mathcal{B}(S^3)$, on note $Pf \in \mathcal{B}(S)$ la fonction $x \mapsto Pf(x) = \int_{S^2} f(x, y, z)P(x, dy, dz)$, quand elle est définie.

Définition 2.2.2 (Chaîne de Markov bifurcantes, voir ([96]))

Soit $(X_n, n \in \mathbb{T})$ une famille de variables aléatoires à valeurs dans S définie sur un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_r, r \in \mathbb{N}), \mathbb{P})$. Soit ν une mesure de probabilité sur (S, \mathcal{S}) et P une \mathbb{T} -probabilité de transition. On dit que $(X_n, n \in \mathbb{T})$ est une (\mathcal{F}_r) -chaîne de Markov bifurcante de mesure initiale ν et de \mathbb{T} -probabilité de transition P si

- X_n est \mathcal{F}_{r_n} -mesurable pour tout $n \in \mathbb{T}$,
- $\mathcal{L}(X_1) = \nu$,
- pour tout $r \in \mathbb{N}$ et pour toute famille $(f_n, n \in \mathbb{G}_r) \subseteq \mathcal{B}_b(S^3)$, on a

$$\mathbb{E} \left[\prod_{n \in \mathbb{G}_r} f_n(X_n, X_{2n}, X_{2n+1}) \middle| \mathcal{F}_r \right] = \prod_{n \in \mathbb{G}_r} Pf_n(X_n).$$

Dans la suite, la filtration implicitement utilisée sera $\mathcal{F}_r = \sigma(X_i, i \in \mathbb{T}_r)$. On note $(Y_r, r \in \mathbb{N})$ la chaîne de Markov sur S avec $Y_0 = X_1$ et de probabilité de transition Q . La chaîne $(Y_r, r \in \mathbb{N})$ correspond à une lignée prise au hasard dans la population.

Pour tout $i \in \mathbb{T}$, on pose $\Delta_i = (X_i, X_{2i}, X_{2i+1})$. On introduit les trois quantités empiriques suivantes

$$(2.3) \quad \overline{M}_{\mathbb{G}_r}(f) = \frac{1}{|\mathbb{G}_r|} \sum_{i \in \mathbb{G}_r} f(\tilde{\Delta}_i), \quad \overline{M}_{\mathbb{T}_r}(f) = \frac{1}{|\mathbb{T}_r|} \sum_{i \in \mathbb{T}_r} f(\tilde{\Delta}_i) \quad \text{et} \quad \overline{M}_n^\Pi(f) = \frac{1}{n} \sum_{i=1}^n f(\tilde{\Delta}_{\Pi(i)}),$$

où $f(\tilde{\Delta}_i) = f(\Delta_i) = f(X_i, X_{2i}, X_{2i+1})$ si $f \in \mathcal{B}(S^3)$ et $f(\tilde{\Delta}_i) = f(X_i)$ si $f \in \mathcal{B}(S)$. Ici Π est une permutation aléatoire qui laisse chaque génération invariante.

Guyon dans [96] a étudié les théorèmes limites des moyennes empiriques (2.3), à savoir la loi des grands nombres (L^2 et presque sûres) et les TCL pour (2.3) quand $f \in \mathcal{B}(S^3)$ mais centré sur l'espérance conditionnelle plutôt que sur la moyenne de la limite.

Les théorèmes limites que nous obtenons, comprennent la loi forte des grands nombres pour la moyenne empirique $\overline{M}_n^\Pi(f)$ avec $f \in \mathcal{B}(S)$ (ce cas n'est pas étudié dans [96]) et la loi du logarithme itéré.

Tous ces résultats sont obtenus sous l'hypothèse d'ergodicité géométrique ou d'ergodicité géométrique uniforme, signifiant que Q^r converge (uniformément) exponentiellement vite vers une mesure limite.

La loi forte des grands nombres est obtenue *via* le contrôle des moments d'ordre 4. Nous généralisons donc le calcul des moments d'ordre 2 fait par Guyon dans [96]. Notons que la technique que nous utilisons peut être appliquée pour calculer d'autres moments d'ordre supérieur mais au prix d'énormes et fastidieux calculs.

Les inégalités de déviations sont obtenues pour des fonctions non bornées, en utilisant l'inégalité de Markov classique et sous l'hypothèse d'ergodicité géométrique. Les résultats sont cependant à ce stade assez restrictifs.

Les inégalités de déviations sont énoncées pour des fonctions bornées et sous l'hypothèse d'ergodicité géométrique uniforme. Leur preuve utilise intensivement l'inégalité d'Azuma-Bennett [11, 19, 100], ce qui nécessite des variables aléatoires bornées. L'extension à des fonctions non bornées et à des conditions d'ergodicité plus faibles peut se faire, en utilisant les inégalités de transport dans l'esprit de [H-D5].

Le principe de déviations modérées est principalement déduit de ces inégalités et des résultats généraux sur les déviations modérées des martingales voir [T-D4]. Leur vitesse dépendra de l'hypothèse satisfaite : ergodicité géométrique uniforme ou seulement ergodicité géométrique.

2.2.1 Cadre des fonctionnelles non-bornées

Soit F un sous espace vectoriel $\mathcal{B}(S)$ tel que

- (i) F contient les constantes,
- (ii) $F^2 := \{f^2/f \in F\} \subset F$,
- (iii) $F \otimes F := \{f \otimes g/f, g \in F\} \subset L^1(P(x, \cdot))$ pour tout $x \in S$, et $P(F \otimes F) := \{Pf/f \in F \otimes F\} \subset F$,
- (iv) il existe une mesure de probabilité μ sur (S, \mathcal{S}) telle que $F \subset L^1(\mu)$ et $\lim_{r \rightarrow \infty} \mathbb{E}_x[f(Y_r)] = (\mu, f)$ pour tout $x \in S$ et $f \in F$,
- (v) pour tout $f \in F$, il existe $g \in F$ telle que pour tout $r \in \mathbb{N}$, $|Q^r f| \leq g$,
- (vi) $F \subset L^1(\nu)$.

Dans ce cadre, on suppose que la chaîne de Markov $(Y_r, r \in \mathbb{N})$ est géométriquement ergodique, c.à.d. :

(H1) Supposons que pour $\forall f \in F$ telle que $(\mu, f) = 0$, $\exists g \in F$ telle que $\forall r \in \mathbb{N}$ et $\forall x \in S$,

$$|Q^r f(x)| \leq \alpha^r g(x) \quad \text{pour un certain } \alpha \in (0, 1).$$

Rappelons que sous cette hypothèse, Guyon [96] a démontré la loi faible des grands nombres pour les trois moyennes empiriques données dans (2.3) et la loi forte des grands nombres uniquement pour $\overline{M}_{\mathbb{G}_r}(f)$,

$\overline{M}_{\mathbb{T}_r}(f)$. Quand $f \in \mathcal{B}(S^3)$ et sous les hypothèses supplémentaires que Pf^2 et Pf^4 existent et appartiennent à F , il a prouvé le TCL pour $\overline{M}_{\mathbb{T}_r}(f)$ et $\overline{M}_n^\Pi(f)$.

Rappelons que le TCL pour les trois moyennes empiriques (2.3) quand $f \in \mathcal{B}(S)$ est une question encore ouverte, voir [60] pour plus de précision.

Dans ce travail, nous complétons les résultats précédents en démontrant la loi forte des grands nombres pour $\overline{M}_n^\Pi(f)$ quand $f \in F$. On prouve aussi la loi du logarithme itéré pour $\overline{M}_n^\Pi(f)$ quand $f \in \mathcal{B}(S^3)$.

Pour établir ces résultats, on a besoin d'un contrôle des moments d'ordre quatre des moyennes empiriques.

Théorème 2.2.3 *Soit $f \in F$ telle que $(\mu, f) = 0$. On a, pour tout $r \in \mathbb{N}$,*

$$(2.4) \quad \mathbb{E} \left[(\overline{M}_{\mathbb{G}_r}(f))^4 \right] \leq \begin{cases} c/4^r & \text{si } \alpha < 1/\sqrt{2} \\ cr^2/4^r & \text{si } \alpha = 1/\sqrt{2} \\ c\alpha^{4r} & \text{si } \alpha > 1/\sqrt{2} \end{cases}$$

où la constante positive $c = c(\alpha, f)$ peut changer d'une ligne à une autre.

Nous obtenons les mêmes estimations pour les moments d'ordre quatre des deux autres moyennes empiriques $\overline{M}_{\mathbb{T}_r}(f)$ et $\overline{M}_n^\Pi(f)$, dans ce dernier cas r est remplacé par r_n dans (2.4). Si $f \in \mathcal{B}(S^3)$ est telle que Pf^2 et Pf^4 existent et appartiennent à F et avec $Pf = 0$, alors pour tout $r \in \mathbb{N}$ et pour une certaine constante positive c , on a $\mathbb{E} \left[(\overline{M}_{\mathbb{G}_r}(f))^4 \right] \leq c/4^r$. On obtient la même estimation pour $\overline{M}_{\mathbb{T}_r}(f)$.

On utilise les notations $M_n^\Pi(f) = \sum_{i=1}^n f(\Delta_{\Pi(i)})$ et $M_{\mathbb{T}_r}(f) = \sum_{i \in \mathbb{T}_r} f(\Delta_i)$. On a le théorème suivant

Théorème 2.2.4

1. *Soit $f \in F$ telle que $(\mu, f) = 0$. On suppose l'hypothèse **(H1)** vérifiée avec $\alpha \in \left(0, \frac{\sqrt[4]{8}}{2}\right)$.*

Alors $\overline{M}_n^\Pi(f)$ converge presque sûrement vers 0.

2. *Soit $f \in \mathcal{B}(S^3)$ telle que $Pf = 0$, Pf^2 et Pf^4 existent et appartiennent à F . Alors*

$$\limsup_{n \rightarrow \infty} \frac{M_n^\Pi(f)}{\sqrt{2 \langle M^\Pi(f) \rangle_n \log \log \langle M^\Pi(f) \rangle_n}} = 1 \quad p.s.$$

$$\text{En particulier} \quad \limsup_{r \rightarrow \infty} \frac{M_{\mathbb{T}_r}(f)}{\sqrt{2 |\mathbb{T}_r| \log \log |\mathbb{T}_r|}} = \sqrt{(\mu, Pf^2)} \quad p.s.$$

$$\text{Si l'hypothèse (H1) est vérifiée avec } \alpha \in \left(0, \frac{\sqrt[4]{8}}{2}\right) \text{ alors } \limsup_{n \rightarrow \infty} \frac{M_n^\Pi(f)}{\sqrt{2n \log \log n}} = \sqrt{(\mu, Pf^2)} \quad p.s.$$

Avec cet énoncé de la première partie sur la loi forte des grands nombres pour $\overline{M}_n^\Pi(f)$, on complète le travail de Guyon [96] où ce résultat est prouvé uniquement pour les deux moyennes empiriques $\overline{M}_{\mathbb{T}_r}(f)$ et $\overline{M}_{\mathbb{G}_r}(f)$. La deuxième partie découle de l'utilisation de la loi du logarithme itérée pour les martingales de Stout [162].

La condition sur α dans ce théorème peut être améliorée, mais le prix à payer pour cela est de faire des calculs énormes liés à l'estimation des moments d'ordres supérieurs. Si f est bornée, ce résultat est vrai pour tout $\alpha \in (0, 1)$ comme on va le voir dans la prochaine partie.

Nous allons donner ensuite quelques inégalités de déviations sous l'hypothèse **(H1)** pour les mesures empiriques (2.3). Cela nous aidera à prouver le résultat des déviations modérées dans ce cadre général où les fonctionnelles des BMC ne sont pas nécessairement bornées.

Théorème 2.2.5 *Pour $[f \in F$ telle que $(\mu, f) = 0]$ ou [pour $f \in \mathcal{B}(S^3)$ telle que Pf et Pf^2 existent et appartiennent à F et $(\mu, Pf) = 0]$, on a pour tout $\delta > 0$ et tout $r \in \mathbb{N}$ et tout $n \in \mathbb{N}$*

$$(2.5) \quad \mathbb{P}\left(|\overline{M}_{G_r}(f)| > \delta\right) \leq \begin{cases} c/(\delta^2 2^r) & \text{si } \alpha < 1/\sqrt{2}, \\ cr/(\delta^2 2^r) & \text{si } \alpha = 1/\sqrt{2}, \\ c\alpha^{2r}/\delta^2 & \text{si } \alpha > 1/\sqrt{2}, \end{cases}$$

où la constante positive $c = c(\alpha, f)$ peut changer d'une ligne à une autre.

Nous obtenons les mêmes estimations pour $\overline{M}_{T_r}(f)$ et pour $\overline{M}_n^\Pi(f)$, dans ce dernier cas r est remplacé par r_n dans (2.5). On introduit la filtration suivante $(\mathcal{H}_n)_{n \geq 0}$ définie par $\mathcal{H}_0 = \sigma(X_1)$ et $\mathcal{H}_n = \sigma(\Delta_{\pi(i)}, 1 \leq i \leq n)$. On peut énoncer maintenant le théorème :

Théorème 2.2.6 *Soit (b_n) une suite croissante de nombres réels positifs telle que*

$$(2.6) \quad \frac{b_n}{\sqrt{n}} \longrightarrow +\infty, \quad \frac{b_n}{\sqrt{n \log n}} \longrightarrow 0, \quad \frac{n}{b_n} \text{ est croissante.}$$

1. *Soit $f \in \mathcal{B}(S^3)$ telle que Pf et Pf^2 existent et appartiennent à F et $(\mu, Pf) = 0$. Alors*

$$\overline{M}_n^\Pi(f) \xrightarrow[\frac{b_n^2}{n}]{\text{superexp}} 0.$$

2. *Soit $f \in \mathcal{B}(S^3)$ telle que Pf^2 et Pf^4 existent et appartiennent à F et que $Pf = 0$. Si*

$$(2.7) \quad \limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \left(n \operatorname{ess\,sup}_{1 \leq k \leq c^{-1}(b_{n+1})} \mathbb{P}\left(|f(\Delta_{\Pi(k)})| > b_n/\mathcal{H}_{k-1}\right) \right) = -\infty,$$

où $c^{-1}(b_{n+1}) := \inf \{k \in \mathbb{N} : \frac{k}{b_k} \geq b_{n+1}\}$, alors $(M_n^\Pi(f)/b_n)$ satisfait le PGD sur \mathbb{R} de vitesse

$$b_n^2/n \text{ et de fonction de taux } I(x) = \frac{x^2}{2(\mu, Pf^2)}.$$

Ce résultat sur les déviations modérées découle des déviations modérées des martingales voir [T-D4], [169, 170]. Une des difficultés principales dans l'application de ce théorème se trouve dans la vérification de la condition (2.7). Notons cependant que dans la gamme de vitesses à considérer, il est suffisant d'avoir un peu de contrôle uniforme sur les X_i , quelques moments de $f(X_i, X_{2i}, X_{2i+1})$ conditionnellement à X_i , qui mène à la condition du type $P|f|^k$ borné pour certains $k \geq 2$. C'est bien sûr le cas si f est bornée. Dans le cas particulier du modèle (2.1), on a pour f telle que $Pf = 0$ et pour tout k

$$\mathbb{E} \left[\exp \left(\lambda \frac{b_n}{n} f(\Delta_{\Pi(k)}) \right) \middle/ \mathcal{H}_{k-1} \right] = \exp \left(\frac{b_n^2}{n} \left(\frac{\lambda^2 P f^2}{2n} \right) (X_{\Pi(k)}) \right).$$

Cette condition implique les déviations modérées pour $(M_n^\Pi(f)/b_n)$.

2.2.2 Cadre des fonctionnelles bornées

Dans ce cadre, on suppose que la chaîne de Markov $(Y_r, r \in \mathbb{N})$ est géométriquement uniformément érgodique, c.à.d. :

(H2) $\exists \mu$ une probabilité sur (S, \mathcal{S}) telle que, $\forall f \in \mathcal{B}_b(S)$ avec $(\mu, f) = 0$, $\exists c$ une constante positive telle que

$$|Q^r f(x)| \leq c\alpha^r \quad \text{pour un certain } \alpha \in (0, 1) \text{ et pour tout } x \in S.$$

Sous l'hypothèse **(H2)**, $\mathcal{B}_b(S)$ vérifie les hypothèses (i)-(vi) de la section précédente. On a

Théorème 2.2.7 Soit $(f \in \mathcal{B}_b(S))$ telle que $(\mu, f) = 0$ ou $(f \in \mathcal{B}_b(S^3))$ telle que $(\mu, Pf) = 0$. Alors on a pour tout $\delta > 0$

$$(2.8) \quad \mathbb{P}(\overline{M}_{\mathbb{G}_r}(f) > \delta) \leq \begin{cases} \exp(c''\delta) \exp(-c'\delta^2|\mathbb{G}_r|), & \text{si } \alpha \leq 1/2, \\ \exp(-c'\delta^2|\mathbb{G}_r|), & \forall r > r_0, \quad \text{si } 1/2 < \alpha < 1/\sqrt{2}, \\ \exp(-c'\delta^2|\mathbb{G}_r|/r), & \forall r > r_0, \quad \text{si } \alpha = 1/\sqrt{2}, \\ \exp(-c'\delta^2/\alpha^{2r}), & \forall r > r_0, \quad \text{si } \alpha > 1/\sqrt{2}, \end{cases}$$

où $r_0 := \log\left(\frac{\delta}{c_0}\right) / \log(\alpha)$, et c_0 , c' et c'' sont des constantes positives qui dépendent de α et f , et qui peuvent changer d'une ligne à une autre.

Nous obtenons le même type d'inégalités de déviations pour $\overline{M}_{\mathbb{T}_r}(f)$ et $\overline{M}_n^\Pi(f)$.

La condition sur α par rapport à $1/2$ est bien sûr liée à la structure binaire de l'arbre. L'extension à l'arbre m -aire suivra les mêmes idées.

Nous pensons qu'il est possible de considérer le cas géométriquement ergodique, et celui des fonctions tests bornées, mais au prix de calculs fastidieux que nous ferons dans un autre travail.

Nous examinerons aussi l'utilisation des inégalités de transport menant à des inégalités de déviations pour des fonctions test Lipschitziennes, sous une certaine propriété de contraction de Wasserstein pour le noyau P , dans l'esprit du théorème 2.5 ou 2.11 dans [H-D5].

Théorème 2.2.8 Soit $f \in \mathcal{B}_b(S^3)$ telle que $(\mu, Pf) = 0$.

1. Alors $\overline{M}_n^\Pi(f)$ converge presque sûrement vers 0.
2. Soit (b_n) une suite croissante de nombres réels positifs telle que $b_n/\sqrt{n} \rightarrow +\infty$, et que
 - si $\alpha < 1/\sqrt{2}$, la suite (b_n) est telle que $b_n/n \rightarrow 0$,
 - si $\alpha = 1/\sqrt{2}$, la suite (b_n) est telle que $(b_n \log n)/n \rightarrow 0$,
 - si $\alpha > 1/\sqrt{2}$, la suite (b_n) est telle que $(b_n \alpha^{r_n+1})/\sqrt{n} \rightarrow 0$.
 Alors $(M_n^\Pi(f)/b_n)$ satisfait le PDM dans S de vitesse b_n^2/n et de fonction de taux $I(x) = x^2/2(\mu, Pf^2)$.

La première partie découle du lemme de Borel-Cantelli et la seconde de l'utilisation du principe de déviations modérées pour les martingales à sauts bornés, voir par exemple [61], [T-D4].

2.2.3 Application : Processus autorégressif bifurcant du premier ordre

Dans cette partie, on applique les résultats sur les chaînes de Markov bifurcantes au processus autorégressif bifurcant d'ordre 1 défini dans (2.1).

Plusieurs extensions du modèle ont été proposées et divers estimateurs sont étudiés dans la littérature pour les paramètres inconnus, voir par exemple [18], [101], [102], [103], [186], [187]. Voir [22] pour des références pertinentes.

On suppose que la distribution ν admet des moments de tous ordres. On a étudié (2.1) dans deux cas :

- le cadre Gaussien qui correspond au cas où $(\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1$ est une suite de variables aléatoires i.i.d. de loi $\mathcal{N}_2(0, \Gamma)$ avec Γ donnée dans (2.2) ;
- le cadre borné qui correspond au cas où X_1 et $(\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1$, est une suite de variables aléatoires i.i.d., prenant ses valeurs dans un ensemble compact. Notons que dans ce cas, $(X_n, n \in \mathbb{T})$ prend ses valeurs dans un ensemble compact.

Notre objectif principal est de donner des inégalités de déviations et le principe de déviations modérées pour l'estimateur des paramètres inconnus $\theta = (\alpha_0, \beta_0, \alpha_1, \beta_1)$ et pour la statistique de test définie dans [96].

Pour estimer $\theta = (\alpha_0, \beta_0, \alpha_1, \beta_1)$, ainsi que σ^2 et ρ , on suppose qu'on observe le sous-arbre complet \mathbb{T}_{r+1} . L'estimateur des moindres carrés $\hat{\theta}^r = (\hat{\alpha}_0^r, \hat{\beta}_0^r, \hat{\alpha}_1^r, \hat{\beta}_1^r)$ de θ est donné par

$$(2.9) \quad \text{pour } \eta \in \{0, 1\} \quad \begin{cases} \hat{\alpha}_\eta^r = \frac{|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i X_{2i+\eta} - \left(|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i \right) \left(|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_{2i+\eta} \right)}{|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i^2 - \left(|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i \right)^2} \\ \hat{\beta}_\eta^r = |\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_{2i+\eta} - \hat{\alpha}_\eta^r |\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i. \end{cases}$$

Notons que dans le cas Gaussien, l'estimateur des moindres carrés correspond à l'estimateur du maximum de vraisemblance. On a aussi besoin d'introduire les estimateurs de la variance conditionnelle σ^2 et la corrélation conditionnelle soeur-soeur ρ . Ces estimateurs sont naturellement donnés par

$$\hat{\sigma}_r^2 = \frac{1}{2|\mathbb{T}_r|} \sum_{i \in \mathbb{T}_r} (\hat{\varepsilon}_{2i}^2 + \hat{\varepsilon}_{2i+1}^2) \quad \hat{\rho}_r = \frac{1}{\hat{\sigma}_r^2} \sum_{i \in \mathbb{T}_r} \hat{\varepsilon}_{2i} \hat{\varepsilon}_{2i+1}$$

où les résidus sont définis par $\hat{\varepsilon}_{2i+\eta} = X_{2i+\eta} - \hat{\alpha}_\eta^r X_i - \hat{\beta}_\eta^r$ pour $\eta \in \{0, 1\}$.

Notons $\mathcal{C}_{pol}^1(\mathbb{R}^3)$ l'ensemble des fonctions $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ qui sont \mathcal{C}^1 et telles que $|f| + |f'|$ est majorée par un polynôme. Alors, on peut montrer que $\mathcal{C}_{pol}^1(\mathbb{R}^3)$ satisfait les hypothèses (i)-(vi). De plus, pour toute $f \in \mathcal{C}_{pol}^1(\mathbb{R}^3)$, l'hypothèse **(H1)** a lieu avec $\alpha = \max(|\alpha_0|, |\alpha_1|)$. Soit μ l'unique distribution stationnaire pour la chaîne de Markov induite $(Y_r, r \in \mathbb{N})$.

Notons \mathbf{x} (resp. \mathbf{x}^2 , \mathbf{xy} , $\mathbf{y} \dots$) les éléments de $\mathcal{C}_{pol}^1(\mathbb{R}^3)$ définis par $(x, y, z) \mapsto x$ (resp. x^2 , xy , y , \dots). On définit deux fonctions continues $\mu_1 : \Theta \rightarrow \mathbb{R}$ et $\mu_2 : \Theta \times \mathbb{R}_+^* \rightarrow \mathbb{R}$ en écrivant $(\mu, \mathbf{x}) = \mu_1(\theta)$ et $(\mu, \mathbf{x}^2) = \mu_2(\theta, \sigma^2)$, où $\theta = (\alpha_0, \beta_0, \alpha_1, \beta_1) \in \Theta = (-1, 1) \times \mathbb{R} \times (-1, 1) \times \mathbb{R}$.

Pour décider entre les hypothèses $H_0 = \{(\alpha_0, \beta_0) = (\alpha_1, \beta_1)\}$ et son alternative $H_1 = \{(\alpha_0, \beta_0) \neq (\alpha_1, \beta_1)\}$, on note $\hat{\mu}_{1,r} = \mu_1(\hat{\theta}^r)$ et $\hat{\mu}_{2,r} = \mu_2(\hat{\theta}^r, \hat{\sigma}_r^2)$, et on utilise la statistique

$$\chi_r^{(1)} = \frac{|\mathbb{T}_r|}{2\hat{\sigma}_r^2(1 - \hat{\rho}_r)} \left\{ (\hat{\alpha}_0^r - \hat{\alpha}_1^r)^2 (\hat{\mu}_{2,r} - \hat{\mu}_{1,r}^2) + \left((\hat{\alpha}_0^r - \hat{\alpha}_1^r) \hat{\mu}_{1,r} + \hat{\beta}_0^r - \hat{\beta}_1^r \right)^2 \right\}.$$

On note $b := \mu_2(\theta, \sigma^2) - \mu_1(\theta)^2$.

Cadre Gaussien Le cadre Gaussien présente des propriétés spécifiques qui permettent des calculs faciles sous des hypothèses plus générales.

Proposition 2.2.9

1. $\forall \delta > 0$, $\forall r \in \mathbb{N}$ et $\forall \gamma < \min\left(\frac{c_1 b}{1+\delta}, \frac{c_1 b}{1+\sqrt{\delta}}, \frac{c_1 b}{1+\sqrt[4]{\delta}}\right)$, où $c_1 := c_1(\mu_1)$ est une constante positive, on a

$$\mathbb{P}\left(\left|\hat{\theta}^r - \theta\right| > \delta\right) \leq \begin{cases} c/(\gamma^{4q} \delta^{4-p} 4^{r+1}) & \text{si } \alpha < 1/\sqrt{2}, \\ c r^2 / (\gamma^{4q} \delta^{4-p} 4^{r+1}) & \text{si } \alpha = 1/\sqrt{2}, \\ c \alpha^{4(r+1)} / (\gamma^{4q} \delta^{4-p}) & \text{si } \alpha > 1/\sqrt{2}, \end{cases}$$

où la constante c dépend α , μ_1 , μ_2 et peut-être différente d'une ligne à une autre, $p = p(\delta) \in \{0, 2, 4\}$ et $q = q(\delta) \in \{0, 1\}$.

2. Soit (b_n) une suite croissante de nombres réels positifs telle que

$$(2.10) \quad \frac{b_n}{\sqrt{n}} \longrightarrow +\infty \quad \text{et} \quad \frac{b_n}{\sqrt{n \log n}} \rightarrow 0.$$

Alors

$$\hat{\theta}^r \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} \theta \quad \text{et} \quad (\hat{\sigma}_r^2, \hat{\rho}_r) \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} (\sigma^2, \rho).$$

3. Soit (b_n) une suite de nombres réels satisfaisant (2.10). Alors la suite $(|\mathbb{T}_r|(\hat{\theta}^r - \theta)/b_{|\mathbb{T}_r|})$ satisfait un PGD sur \mathbb{R}^4 de vitesse $b_{|\mathbb{T}_r|}^2/|\mathbb{T}_r|$ et de fonction de taux I donnée par $I(x) = \frac{1}{2}x^t(\Sigma')^{-1}x$, où

$$\Sigma' = \sigma^2 \begin{pmatrix} K & \rho K \\ \rho K & K \end{pmatrix} \quad \text{avec} \quad K = \frac{1}{\mu_2(\theta, \sigma^2) - \mu_1(\theta)^2} \begin{pmatrix} 1 & -\mu_1(\theta) \\ -\mu_1(\theta) & \mu_2(\theta, \sigma^2) \end{pmatrix}.$$

4. Soit (b_n) une suite de nombres réels satisfaisant (2.10). Alors sous l'hypothèse nulle $H_0 = \{(\alpha_0, \beta_0) = (\alpha_1, \beta_1)\}$, $|\mathbb{T}_r|^{1/2}(\chi_r^{(1)})^{1/2}/b_{|\mathbb{T}_r|}$ satisfait un PGD sur \mathbb{R} de vitesse $b_{|\mathbb{T}_r|}^2/|\mathbb{T}_r|$ et de fonction de taux

$$I'(y) = \begin{cases} y^2/2 & \text{si } y \in \mathbb{R}_+ \\ +\infty & \text{ailleurs.} \end{cases}$$

Sous l'hypothèse alternative H_1 de H_0 , on a $\forall A > 0 \quad \limsup_{r \rightarrow \infty} \frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}^2} \log \mathbb{P}(\chi_r^{(1)} < A) = -\infty$.

Les valeurs de p et q dépendent de l'ordre de δ . Par exemple, si δ est assez petit, on a $p = 0$ et $q = 0$.

Cadre compact : le cas uniformément ergodique. Nous supposons ici que le bruit et l'état initial X_1 prennent leurs valeurs dans un ensemble compact. La nouveauté est que la gamme de vitesses est améliorée par rapport au cadre Gaussien.

On prend $F = \mathcal{C}_b^1(\mathbb{R})$ l'ensemble de toutes les fonctions f dans \mathcal{C}^1 bornées sur \mathbb{R} . Donc, on peut facilement montrer que l'hypothèse **(H2)** est vérifiée avec $\alpha = \max(|\alpha_0|, |\alpha_1|)$.

Pour (b_n) une suite de nombres réels satisfaisant les hypothèses du théorème 2.2.8 (au lieu de (2.10)), les résultats limites du cadre Gaussien restent valables. Notons que pour la preuve des déviations modérées dans ce cadre, on n'a pas besoin d'utiliser la méthode des cumulants, contrairement au cadre Gaussien. Grâce aux déviations modérées des martingales à sauts bornés (voir [61]), on a seulement besoin de prouver la convergence superexponentielle du processus croissant associé à la martingale.

Nous pouvons maintenant affiner les résultats limites en donnant l'inégalité de déviations suivante :

Proposition 2.2.10

$\forall \delta > 0$ et $\forall \gamma < \min\left(\frac{c_1 b}{1+\delta}, \frac{c_1 b}{1+\sqrt{\delta}}, \frac{c_1 b}{1+\sqrt[3]{\delta}}\right)$, où $c_1 := c_1(\mu_1)$ est une constante positive, avec $\kappa(\delta) = \gamma^q \delta^{1-p/2}$, on a

$$\mathbb{P}\left(\left|\hat{\theta}^r - \theta\right| > \delta\right) \leq \begin{cases} c_2 \exp(c'' \kappa(\delta) - c' \kappa(\delta)^2 |\mathbb{T}_r|), & \forall r \in \mathbb{N}, & \text{si } \alpha < 1/2 \\ c_2 \exp(c'(r+1)\kappa(\delta) - c'' \kappa(\delta)^2 |\mathbb{T}_r|), & \forall r \in \mathbb{N}, & \text{si } \alpha = 1/2 \\ c_2 \exp(-c' \kappa(\delta)^2 |\mathbb{T}_r|), & \forall r > r_0, & \text{si } 1/2 < \alpha < 1/\sqrt{2} \\ c_2 \exp(-c' \kappa(\delta)^2 |\mathbb{T}_r| / ((r+1)\gamma)), & \forall r > r_0, & \text{si } \alpha = 1/\sqrt{2} \\ c_2 \exp(-c' \kappa(\delta)^2 / \alpha^{2(r+1)}), & \forall r > r_0, & \text{si } \alpha > 1/\sqrt{2}, \end{cases}$$

où $r_0 := \log(\gamma^q \delta^{1-p/2}/c_0) / \log \alpha$, c_2 est une constante positive, c' et c'' dépendent de α et c , c_0 dépend de α , c et γ . Ces constantes peuvent changer d'une ligne à une autre, $p \in \{0, 1, 3/2\}$ et $q \in \{0, 1\}$.

2.3 Processus auto-régressif bifurcant d'ordre p

Soit $p \in \mathbb{N}^*$. On considère le processus assymétrique $\text{BAR}(p)$ donné, pour $n \geq 2^{p-1}$, par

$$\begin{cases} X_{2n} = a_0 + \sum_{k=1}^p a_k X_{[\frac{n}{2^k-1}]} + \varepsilon_{2n} \\ X_{2n+1} = b_0 + \sum_{k=1}^p b_k X_{[\frac{n}{2^k-1}]} + \varepsilon_{2n+1}, \end{cases}$$

où la notation $[x]$ désigne la partie entière x . Il a été introduit par Bercu et al. [22]. L'état initial $\{X_k, 1 \leq k \leq 2^{p-1} - 1\}$ sont les ancêtres tandis que $(\varepsilon_{2n}, \varepsilon_{2n+1})$ est le bruit du processus. Les paramètres (a_0, a_1, \dots, a_p) et (b_0, b_1, \dots, b_p) sont des vecteurs réels inconnus.

Comme pour les processus autorégressifs standards, on peut réécrire ce modèle comme un $\text{BAR}(1)$ d'ordre multidimensionnel pour tout $n \geq 2^{p-1}$ sous la forme

$$\begin{cases} \mathbb{X}_{2n} = A\mathbb{X}_n + \eta_{2n} \\ \mathbb{X}_{2n+1} = B\mathbb{X}_n + \eta_{2n+1} \end{cases}$$

où $\mathbb{X}_n = \left(X_n, X_{[\frac{n}{2}]}, \dots, X_{[\frac{n}{2^{p-1}-1}]}\right)^T$ est le vecteur de régression, $\eta_{2n} = (a_0 + \varepsilon_{2n})e_1$ et $\eta_{2n+1} = (b_0 + \varepsilon_{2n+1})e_1$, avec $e_1 = (1, 0, \dots, 0)^t \in \mathbb{R}^p$. Les matrices associées A et B d'ordre $p \times p$ sont données par

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_p \\ 1 & 0 & \cdots & 0 \\ 0 & \cdot & \cdot & \cdot \\ 0 & \cdot & 1 & \cdot \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_1 & b_2 & \cdots & b_p \\ 1 & 0 & \cdots & 0 \\ 0 & \cdot & \cdot & \cdot \\ 0 & \cdot & 1 & \cdot \end{pmatrix}.$$

Ce processus est une généralisation directe du $\text{BAR}(p)$ symétrique étudié par Huggins, Bassawa et Zhou. Dans le cas particulier où $p = 1$, il correspond au modèle étudié par Guyon. Ici la principale difficulté à gérer est le caractère multidimensionnel des différentes variables. On suppose que les matrices A et B vérifient la condition de contractivité suivante

$$(2.11) \quad \beta = \max(\|A\|, \|B\|) < 1.$$

Notons θ le vecteur des paramètres du modèle et $\hat{\theta}_n$ son estimateur, définis par

$$\theta = \begin{pmatrix} a_0 \\ \cdot \\ \cdot \\ a_p \\ b_0 \\ \cdot \\ \cdot \\ b_p \end{pmatrix} \quad \text{et} \quad \hat{\theta}_n = \begin{pmatrix} \hat{a}_{0,n} \\ \cdot \\ \cdot \\ \hat{a}_{p,n} \\ \hat{b}_{0,n} \\ \cdot \\ \cdot \\ \hat{b}_{p,n} \end{pmatrix}.$$

On note $\Sigma_n = I_2 \otimes S_n$, où \otimes signifie le produit matriciel de Kronecker et la matrice S_n d'ordre $(p+1) \times (p+1)$ est définie comme suit

$$S_n = \sum_{k \in \mathbb{T}_{n,p-1}} \begin{pmatrix} 1 & \mathbb{X}_k^t \\ \mathbb{X}_k & \mathbb{X}_k \mathbb{X}_k^t \end{pmatrix}.$$

Comme dans Bercu et al. [22], on introduit l'estimateur des moindres carrés $\hat{\theta}_n$ de θ pour tout $n \geq p$, à partir des observations de tous les individus jusqu'à la n -ième génération (c.à.d. le sous-arbre complet \mathbb{T}_n). Pour $n \geq p$, cet estimateur est donné par

$$\hat{\theta}_n = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1,p-1}} \begin{pmatrix} X_{2k} \\ X_k \mathbb{X}_{2k} \\ X_{2k+1} \\ X_k \mathbb{X}_{2k+1} \end{pmatrix},$$

où $\mathbb{T}_{n,p} = \{k \in \mathbb{T}_n, k \geq 2^p\}$ est le sous-arbre de tous les individus entre la p -ième et la n -ième génération (\mathbb{T}_{p-1} est exclu). Par conséquent, on obtient

$$(2.12) \quad \hat{\theta}_n - \theta = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1,p-1}} \begin{pmatrix} \varepsilon_{2k} \\ \varepsilon_{2k} \mathbb{X}_k \\ \varepsilon_{2k+1} \\ \varepsilon_{2k+1} \mathbb{X}_k \end{pmatrix}.$$

On désigne par $\mathbb{F} = (\mathcal{F}_n)$ la filtration naturelle associée au processus BAR(p). La tribu \mathcal{F}_n est donc la σ -algèbre engendrée par les individus jusqu'à la n -ième génération.

Pour les états initiaux, on pose $\overline{X}_1 = \max \left\{ \|\mathbb{X}_k\|, k \leq 2^{p-1} \right\}$ avec la convention que $X_0 = 0$ et on introduit l'hypothèse suivante :

(Xa) Pour un certain a , il existe $\zeta > 0$ tel que $\mathbb{E} \left[\exp \left(\zeta \overline{X}_1^a \right) \right] < \infty$.

Pour le bruit $(\varepsilon_{2n}, \varepsilon_{2n+1})$, les hypothèses peuvent être de deux types.

1. Dans le premier cas, on suppose l'indépendance du bruit, ce qui permet d'imposer des conditions moins restrictives sur l'intégrabilité exponentielle du bruit.

Cas 1 : On suppose que $((\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1)$ est une suite de variables aléatoires i.i.d. centrées de matrice de covariance Γ donnée par

$$(2.13) \quad \Gamma = \begin{pmatrix} \sigma^2 & \rho \\ \rho & \sigma^2 \end{pmatrix}, \quad \text{où } \sigma^2 > 0 \text{ et } |\rho| < \sigma^2.$$

Pour tout $n \geq p-1$ et tout $k \in \mathbb{G}_n$, on pose

$$\mathbb{E}[\varepsilon_k^2] = \sigma^2, \quad \mathbb{E}[\varepsilon_k^4] = \tau^4, \quad \mathbb{E}[\varepsilon_{2k}\varepsilon_{2k+1}] = \rho, \quad \mathbb{E}[\varepsilon_{2k}^2\varepsilon_{2k+1}^2] = \nu^2 \quad \text{où } \tau^4 > 0 \quad \text{et} \quad \nu^2 < \tau^4.$$

On suppose en plus que la condition (X2) sur l'état initial est satisfaite et que

(G2) on peut trouver $\gamma > 0$ et $c > 0$ tels que pour tout $n \geq p-1$, pour tout $k \in \mathbb{G}_n$ et pour tout t tel que $|t| \leq c$, on ait

$$\mathbb{E} \left[\exp \left(t (\varepsilon_k^2 - \sigma^2) \right) \right] \leq \exp \left(\frac{\gamma t^2}{2} \right).$$

Dans ce cas, on impose la condition supplémentaire suivante sur la vitesse des déviations

(V1) la vitesse de déviations (v_n) est une suite croissante de nombres réels telle que $v_n \rightarrow +\infty$ et pour β donné par (2.11),

- si $\beta \leq 1/2$, la suite (v_n) est telle que $\frac{v_n \log n}{\sqrt{n}} \rightarrow 0$,
- si $\beta > 1/2$, la suite (v_n) est telle que $(v_n \sqrt{\log n}) \beta^{\frac{r_n+1}{2}} \rightarrow 0$.

2. Dans le deuxième cas, contrairement au premier, on ne suppose plus que la suite $((\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1)$ est i.i.d.. Le prix à payer pour l'abandon de cette hypothèse est de supposer des conditions plus fortes d'intégrabilité des moments exponentiels. En effet, on a besoin de cela pour utiliser le principe de déviations modérées pour les martingales, spécialement pour pouvoir la condition de Lindeberg en utilisant la condition de Lyapunov.

Cas 2 : On suppose que pour tout $n \geq p-1$ et tout $j \in \mathbb{G}_{n+1}$ $\mathbb{E}[\varepsilon_j/\mathbb{F}_n] = 0$ et pour tout $k, l \in \mathbb{G}_{n+1}$ différents et tels que $[\frac{k}{2}] \neq [\frac{l}{2}]$, ε_k et ε_l sont indépendants conditionnellement à \mathcal{F}_n . On utilise les mêmes notations que dans le cas 1 : pour tout $n \geq p-1$ et tout $k \in \mathbb{G}_{n+1}$,

$$\mathbb{E}[\varepsilon_k^2/\mathcal{F}_n] = \sigma^2, \quad \mathbb{E}[\varepsilon_k^4/\mathcal{F}_n] = \tau^4, \quad \mathbb{E}[\varepsilon_{2k}\varepsilon_{2k+1}/\mathcal{F}_n] = \rho, \quad \mathbb{E}[\varepsilon_{2k}^2\varepsilon_{2k+1}^2/\mathcal{F}_n] = \nu^2 \quad p.s.$$

où $\tau^4 > 0$, $\nu^2 < \tau^4$ et on utilise aussi la notation Γ pour désigner la matrice des covariances conditionnelles associées à $(\varepsilon_{2n}, \varepsilon_{2n+1})$. Dans ce cas, on suppose en plus que la condition (Xa) pour un $a > 2$ sur l'état initial est satisfaite, et que les conditions suivantes sont satisfaites :

(Ea) $\exists a > 2, \exists t > 0$ et $\exists E > 0$ tels que $\forall n \geq p-1$ et $\forall k \in \mathbb{G}_{n+1}$, $\mathbb{E}[\exp(t|\varepsilon_k|^{2a})/\mathcal{F}_n] \leq E$ p.s.

Dans ce cas, on introduit les hypothèses suivantes sur la vitesse des déviations

(V2) (v_n) est une suite croissante de nombres réels positifs telle que $v_n \rightarrow +\infty$, et pour β donné par (2.11),

- si $\beta < 1/\sqrt{2}$, la suite (v_n) est telle que $\frac{v_n \log n}{\sqrt{n}} \rightarrow 0$,
- si $\beta = 1/\sqrt{2}$, la suite (v_n) est telle que $\frac{v_n (\log n)^{3/2}}{\sqrt{n}} \rightarrow 0$,
- si $\beta > 1/\sqrt{2}$, la suite (v_n) est telle que $(v_n \log n)\beta^{r_n+1} \rightarrow 0$.

La condition sur la vitesse des déviations dans le cas 2 est plus restrictive que celle dans le cas 1, car dans ce dernier cas on suppose une condition d'intégrabilité **(Ea)** plus forte sur le bruit. Cette condition sur la vitesse des déviations apparaît naturellement dans les calculs. Plus précisément, le terme en log provient de l'interversion de la probabilité et de la somme. De [H-D6] ou [115], on déduit de la condition **(Ea)** que

(N1) $\exists \phi > 0$ tel que $\forall n \geq p-1, \forall k \in \mathbb{G}_{n+1}$ et $\forall t \in \mathbb{R}$, $\mathbb{E}[\exp(t\varepsilon_k)/\mathcal{F}_n] < \exp(\phi t^2/2)$, p.s.

Dans le premier cas, on obtient la même conclusion sans le conditionnement, c.à.d.

(G1) $\exists \phi > 0$ tel que $\forall n \geq p-1, \forall k \in \mathbb{G}_n$ et $\forall t \in \mathbb{R}$, $\mathbb{E}[\exp(t\varepsilon_k)] < \exp(\phi t^2/2)$.

Armé des récents développements de la théorie des inégalités de transport, de l'intégrabilité exponentielle et des inégalités fonctionnelles (voir Ledoux [115], Gozlan [90] et Gozlan et Leonard [91]), on peut prouver qu'une condition suffisante pour que l'hypothèse **(G2)** soit satisfaite est l'existence de $t_0 > 0$ tel que, pour tout $n \geq p-1$ et pour tout $k \in \mathbb{G}_n$, on ait $\mathbb{E}[\exp(t_0 \varepsilon_k^2)] < \infty$.

Avant d'énoncer nos principaux résultats, fixons encore quelques notations. On introduit

$$\bar{a} = \frac{a_0 + b_0}{2}, \quad \bar{a}^2 = \frac{a_0^2 + b_0^2}{2}, \quad \bar{A} = \frac{A + B}{2}, \quad \Xi = \bar{a}(I_p - \bar{A})^{-1}e_1, \quad L = \begin{pmatrix} 1 & \Xi \\ \Xi & \Lambda \end{pmatrix} \quad \text{et} \quad \Sigma = I_2 \otimes L,$$

où Λ est l'unique solution de l'équation suivante (voir Lemma A.4 dans [22]) :

$$\Lambda = T + \frac{1}{2}(A\Lambda A^t + B\Lambda B^t) \quad \text{avec} \quad T = \left(\sigma^2 + \bar{a}^2\right) e_1 e_1^t + \frac{1}{2} \left(a_0 (A\Xi e_1^t + e_1 \Xi^t A^t) + b_0 (B\Xi e_1^t + e_1 \Xi^t B^t)\right).$$

2.3.1 Principaux résultats

Dans ce premier théorème, nous donnons les inégalités de déviations pour l'estimateur des paramètres.

Théorème 2.3.1 *Pour tout $\delta > 0$ et tout $\ell > 0$ tels que $\ell < \|\Sigma\|/(1+\delta)$, on a*

(i) *dans le cas 1, en notant $c(\delta, \ell) := c_2(\delta\ell)^2/(c_3 + \delta\ell)$,*

$$\mathbb{P}\left(|\hat{\theta}_n - \theta| > \delta\right) \leq \begin{cases} c_1 \exp(-2^n c(\delta, \ell)/(n-1)^2) & \text{si } \beta < 1/2, \\ c_1(n-1) \exp(-2^n c(\delta, \ell)/(n-1)^2) & \text{si } \beta = 1/2, \\ c_1(n-1) \exp(-c(\delta, \ell)/((n-1)\beta^n)) & \text{si } \beta > 1/2, \end{cases}$$

(ii) *dans le cas 2, en notant $\kappa(\delta, \ell) := c_2(\delta\ell)^2/(c_3 + c_4\delta\ell)$*

$$\mathbb{P}\left(|\hat{\theta}_n - \theta| > \delta\right) \leq \begin{cases} c_1 \exp(-2^n \kappa(\delta, \ell)/(n-1)^2) & \text{si } \beta < 1/\sqrt{2}, \\ c_1 \exp(-2^n \kappa(\delta, \ell)/(n-1)^3) & \text{si } \beta = 1/\sqrt{2}, \\ c_1 \exp(-\kappa(\delta, \ell)/((n-1)^2 \beta^{2n})) & \text{si } \beta > 1/\sqrt{2}. \end{cases}$$

Dans les deux cas, les constantes c_1, c_2 et c_3 dépendent de σ^2, β, γ et ϕ et peuvent changer d'une ligne à une autre et sont telles que $c_1, c_2 > 0, c_3, c_4 \geq 0, (c_3, c_4) \neq (0, 0)$.

On peut souligner que des calculs ennuyeux mais directs permettent d'estimer les constantes qui apparaissent dans le théorème précédent.

Les majorations dans le théorème précédent sont obtenues pour un $n \geq p-1$ arbitraire (ce n'est donc pas un résultat limite, contrairement aux résultats ci-dessous). Par conséquent, elles sont très pratiques dans la statistique non asymptotique quand la taille de l'échantillon ne permet pas l'utilisation de théorèmes limites.

Dans le résultat suivant, on présente les déviations modérées de l'estimateur $\hat{\theta}_n$.

Théorème 2.3.2 *Dans le cas 1 ou le cas 2, la suite $\left(\sqrt{|\mathbb{T}_{n-1}|}(\hat{\theta}_n - \theta)/v_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$ satisfait le PGD sur $\mathbb{R}^{2(p+1)}$ de vitesse $v_{|\mathbb{T}_{n-1}|}^2$ et de fonction de taux*

$$I_\theta(x) = \sup_{\lambda \in \mathbb{R}^{2(p+1)}} \{\lambda^t x - \lambda(\Gamma \otimes L^{-1})\lambda^t\} = \frac{1}{2}x^t(\Gamma \otimes L^{-1})^{-1}x.$$

Pour prouver ce théorème, on utilise l'approche martingale. Pour tout $n \geq p$, on pose

$$M_n = \sum_{k \in \mathbb{T}_{n-1, p-1}} \begin{pmatrix} \varepsilon_{2k} \\ \varepsilon_{2k} \mathbb{X}_k \\ \varepsilon_{2k+1} \\ \varepsilon_{2k+1} \mathbb{X}_k \end{pmatrix} \in \mathbb{R}^{2(p+1)}.$$

On peut clairement réécrire (2.12) comme $\hat{\theta}_n - \theta = \Sigma_{n-1}^{-1} M_n$. On sait grâce à Bercu et al. [22] que (M_n) est une martingale de carré intégrable adaptée à la filtration $\mathcal{F} = (\mathcal{F}_n)$. Son processus croissant associé est donné, pour tout $n \geq p$, par $\langle M \rangle_n = \Gamma \otimes S_{n-1}$, où S_n est donnée dans (2.3) et Γ est donnée dans (2.13).

La stratégie de la preuve des déviations modérées consiste en les étapes suivantes :

- obtenir la convergence super-exponentielle de la variation quadratique de la martingale (M_n) . Cette étape est cruciale et c'est la clé de la preuve du reste des résultats. Elle est réalisée au moyen d'inégalités exponentielles puissantes. Ceci nous permet d'obtenir les inégalités de déviations pour l'estimateur des paramètres,
- introduire une martingale tronquée qui satisfait les déviations modérées, grâce à [T-D4],
- vérifier que la martingale tronquée est une bonne approximation exponentielle de (M_n) , au sens des déviations modérées.

Nous allons maintenant considérer les estimateurs des paramètres dans le processus de bruit. D'un côté, nous proposons d'estimer la variance conditionnelle σ^2 par :

$$\hat{\sigma}_n^2 = \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} \|\hat{V}_k\|^2 = \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} (\hat{\varepsilon}_{2k}^2 + \hat{\varepsilon}_{2k+1}^2),$$

où, pour tout $n \geq p-1$ et tout $k \in \mathbb{G}_n$, on a $\hat{V}_k^T = (\hat{\varepsilon}_{2k}, \hat{\varepsilon}_{2k+1})^T$ avec

$$\begin{cases} \hat{\varepsilon}_{2k} = X_{2k} - \hat{a}_{0,n} - \sum_{i=1}^p \hat{a}_{i,n} X_{[\frac{k-i}{2^{i-1}}]} \\ \hat{\varepsilon}_{2k+1} = X_{2k+1} - \hat{b}_{0,n} - \sum_{i=1}^p \hat{b}_{i,n} X_{[\frac{k-i}{2^{i-1}}]} \end{cases}.$$

D'un autre côté, on estime la covariance conditionnelle ρ par $\hat{\rho}_n = \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} \hat{\varepsilon}_{2k} \hat{\varepsilon}_{2k+1}$.

On introduit aussi $\sigma_n^2 = \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p}} (\varepsilon_{2k}^2 + \varepsilon_{2k+1}^2)$ et $\rho_n = \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p}} \varepsilon_{2k} \varepsilon_{2k+1}$.

Théorème 2.3.3 *Soit (v_n) une suite croissante de nombres réels positifs telle que $v_n \rightarrow \infty$ et $\frac{v_n}{\sqrt{n}} \rightarrow 0$. Dans le cas 1 ou 2,*

- (1) la suite $\left(\sqrt{|\mathbb{T}_{n-1}|}(\sigma_n^2 - \sigma^2)/v_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$ satisfait le PGD sur \mathbb{R} de vitesse $v_{|\mathbb{T}_{n-1}|}^2$ et de fonction de taux $I_{\sigma^2}(x) = \frac{x^2}{\tau^4 - 2\sigma^4 + \nu^2}$.
- (2) la suite $\left(\sqrt{|\mathbb{T}_{n-1}|}(\rho_n - \rho)/v_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$ satisfait le PGD sur \mathbb{R} de vitesse $v_{|\mathbb{T}_{n-1}|}^2$ et de fonction de taux $I_{\rho}(x) = \frac{x^2}{2(\nu^2 - \rho^2)}$.

Nous remarquons qu'il n'y a aucune restriction sur la vitesse (v_n) ici. Il est plus intéressant de prouver les déviations modérées avec $\hat{\sigma}_n^2$ à la place de σ_n^2 , ce qui est le cas si on peut prouver par exemple que $\left(\sqrt{|\mathbb{T}_{n-1}|}(\hat{\sigma}_n^2 - \sigma^2)/v_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$ et $\left(\sqrt{|\mathbb{T}_{n-1}|}(\sigma_n^2 - \sigma^2)/v_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$ sont exponentiellement équivalents au sens des grandes déviations. On aurait donc besoin de la convergence suivante : $\frac{\sqrt{|\mathbb{T}_{n-1}|}}{v_{|\mathbb{T}_{n-1}|}}(\hat{\sigma}_n^2 - \sigma_n^2) \xrightarrow[v_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0$.

0. La preuve est très technique et très restrictive pour la vitesse (v_n) des déviations modérées. Actuellement, on peut seulement prouver que $\hat{\sigma}_n^2 - \sigma_n^2 \xrightarrow[v_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0$.

Dans le théorème suivant, on donne ainsi quelques convergences superexponentielles.

Théorème 2.3.4 Dans les deux cas, on a $\hat{\sigma}_n^2 \xrightarrow[v_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} \sigma^2$.

Dans le cas 1, si on remplace **(G2)**, par

(G2') on peut trouver $\gamma' > 0$ tel que, pour tout $n \geq p-1$, pour tous $k, l \in \mathbb{G}_{n+1}$ tels que $\lfloor \frac{k}{2} \rfloor = \lfloor \frac{l}{2} \rfloor$ et pour tous $t \in]-c, c[$ pour un certain $c > 0$, on ait $\mathbb{E}[\exp t(\varepsilon_k \varepsilon_l - \rho)] \leq \exp\left(\frac{\gamma' t^2}{2}\right)$,

et dans le cas 2, si on remplace **(Ea)** par l'hypothèse suivante

(E2') on peut trouver $\gamma' > 0$ tel que, pour tous $n \geq p-1$, pour tous $k, l \in \mathbb{G}_{n+1}$ tels que $\lfloor \frac{k}{2} \rfloor = \lfloor \frac{l}{2} \rfloor$ et pour tous $t \in \mathbb{R}$, on ait

$$\mathbb{E}[\exp t(\varepsilon_k \varepsilon_l - \rho) / \mathcal{F}_n] \leq \exp\left(\frac{\gamma' t^2}{2}\right), \quad p.s.$$

alors dans le cas 1 ou 2, on a $\hat{\rho}_n \xrightarrow[v_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} \rho$.

Problèmes et perspectives

1. Il nous paraît nécessaire de poursuivre l'étude des chaînes de Markov bifurcantes pour obtenir un principe de grandes déviations. On souhaite obtenir la borne supérieure et la borne inférieure des grandes déviations de la mesure empirique suivante :

$$L_n = \frac{1}{|\mathbb{T}_n|} \sum_{i \in \mathbb{T}_n} \delta_{(X_i, X_{2i}, X_{2i+1})}.$$

En s'inspirant du travail bien connu de Donsker-Varadhan, on peut obtenir la borne inférieure par un changement de probabilité et une loi des grands nombres. Il reste l'investigation de la borne supérieure et la comparaison des fonctions de taux des deux bornes.

2. Les chaînes de Markov bifurcantes sur un arbre de Galton-Watson ont été introduites pour modéliser la généalogie incomplète et prendre en compte la mort possible des cellules. Nous souhaitons continuer l'étude des inégalités de déviations et du transport pour ce cadre, voir [56], [57], [60].
3. L'étude des processus autorégressifs bifurcants à coefficients aléatoires a permis d'étendre les processus autorégressifs bifurcants en généralisant le côté aléatoire de l'évolution, (voir par exemple [20], [21], [28]). Le processus BAR à coefficients aléatoires est défini par la donnée d'un ancêtre initial X_1 , de deux suites de bruits $(\varepsilon_{2k}, \varepsilon_{2k+1})_{k \geq 1}$ et $(\eta_{2k}, \eta_{2k+1})_{k \geq 1}$ et par l'équation récursive $X_{2k+\delta} = c + (d + \eta_{2k+\delta})X_k + \varepsilon_{2k+\delta}$, avec $\delta \in \{0, 1\}$.

On peut le réécrire sous une forme similaire au BAR(1). Mais dans ce cas, on retrouve l'autorégression à l'intérieur du bruit ce qui pose des problèmes supplémentaires et impose un changement d'approche par rapport au traitement du processus BAR standard à coefficients fixes.

Le modèle de processus autorégressif bifurquant asymétrique avec des coefficients aléatoires couplé avec un arbre Galton-Watson pour prendre en compte des observations probablement manquantes a aussi été considéré. Pour les théorèmes limites classiques, voir [58]. Nous souhaitons obtenir quelques inégalités de déviations pour ces modèles.

4. Le TCL pour $\sum_{i \in \mathbb{T}_n} g(X_i)$ reste une question ouverte. L'approche classique basée sur l'utilisation de l'équation de Poisson $G - PG = g - \mu(g)$ pour étudier les fluctuations ne marche pas (μ étant la mesure invariante du noyau de transition P). En effet, avec la décomposition

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (g(X_i) - \mu(g)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (G(X_i) - PG(X_{i-1})) + \frac{1}{\sqrt{n}} PG(X_0) - \frac{1}{\sqrt{n}} PG(X_n)$$

et en utilisant la théorie des martingales, on obtient le TCL pour le premier terme dans le membre droite de l'égalité précédente. Il reste juste à montrer que les deux derniers termes convergent vers 0 pour conclure.

Mais dans notre cas, le dernier terme de bord $\frac{1}{\sqrt{n}} PG(X_n)$ devient le terme de bord sur la dernière génération \mathbb{G}_n dont le cardinal est du même ordre que $|\mathbb{T}_n|$. Donc l'ordre du terme de bord n'est pas négligeable, ce qui nous empêche de conclure.

Chapitre 3

Inégalités fonctionnelles et applications

Dans ce chapitre, nous présentons brièvement les résultats des articles [H-D5] publié dans *Annals of Probability* et [H-D7] publié dans *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*.

3.1 Introduction

Sur (E, d) un espace métrique, pour $p \geq 1$ et μ, ν deux mesures de probabilités sur E , on définit la distance de Wasserstein L^p

$$(3.1) \quad W_p^d(\mu, \nu) = \inf \left(\int \int d(x, y)^p d\pi(x, y) \right)^{1/p}$$

où l'infimum est pris sur toutes les mesures de probabilités sur $E \times E$ de marginales μ et ν (π est alors appelée un couplage de μ et ν). On dit alors que μ satisfait une inégalité de transport T_p , pour $p \in [1, 2]$, de constante C (et on note $\mu \in T_p(C)$), si pour toute mesure de probabilité ν , on a

$$(3.2) \quad W_p^d(\mu, \nu) \leq \sqrt{2CH(\nu|\mu)}$$

avec $H(\nu|\mu)$ l'entropie (ou information de Kullback) définie par

$$H(\nu|\mu) = \begin{cases} \int \log \frac{d\nu}{d\mu} d\nu, & \text{si } d\nu \ll d\mu, \\ +\infty & \text{sinon.} \end{cases}$$

Il est intéressant de noter que la célèbre inégalité de Csiszar-Kullback-Pinsker rentre dans le cadre de cette définition. Signalons deux classes particulières d'inégalités de transport apparaissant fréquemment dans la littérature : $T_1(C)$ et $T_2(C)$. L'inégalité de Jensen montre immédiatement que $T_2(C)$ entraîne $T_1(C)$.

L'inégalité T_2 est nettement plus structurée que l'inégalité T_1 . Tout d'abord, comme l'ont montré Otto et Villani dans [OV00], l'inégalité T_2 entraîne une inégalité de trou spectral, également appelée inégalité de Poincaré : $\mu \in TS(C)$ si pour toute fonction f régulière

$$(3.3) \quad \text{Var}_\mu(f) \leq C \int |\nabla f|^2 d\mu.$$

L'importance du trou spectral réside dans le fait qu'il décrit le taux de la convergence exponentielle. L'étude de la constante du trou spectral a une très longue histoire, le lecteur est renvoyé aux livres [44, 168] pour des bibliographies.

L'étude des inégalités de transport a réellement été lancée par les travaux de Marton [125, 124, 126] et Talagrand [163] dans les années 90. La motivation principale de ces travaux et de ceux qui ont suivi était les diverses inégalités de concentration de la mesure qui peuvent être tirées des inégalités de type (3.2). A ce titre, les premiers travaux de Marton portent sur des raffinements multidimensionnels de l'inégalité de Csiszar-Kullback-Pinsker qui permettent de retrouver certaines inégalités de concentration universelles pour les mesures produits obtenues par Talagrand. Dans l'article [163], Talagrand a introduit quant à lui l'inégalité de transport T_2 et a prouvé que la mesure Gaussienne standard sur \mathbb{R} vérifie $T_2(C)$ avec une constante optimale.

Même si l'inégalité T_2 est en général plus délicate à montrer, les critères connus étant en général difficiles à vérifier, la grande force de celle-ci est qu'elle se tensorise facilement. On peut ainsi obtenir de bonnes estimations en grande dimension sans avoir à payer le prix d'une trop grosse dégradation de la constante.

Le résultat fondateur dans le domaine est la démonstration par Talagrand du fait que la mesure Gaussienne standard sur \mathbb{R}^n , muni de la distance euclidienne, vérifie l'inégalité $T_2(2)$ avec en particulier une constante qui ne dépend pas de la dimension n de l'espace.

Définissons l'inégalité de Sobolev logarithmique : on dira que $\mu \in ISL(C)$ si pour toute fonction f régulière

$$(3.4) \quad \text{Ent}_\mu(f^2) := \int f^2 \log \left(\frac{f^2}{\int f^2 d\mu} \right) d\mu \leq C \int |\nabla f|^2 d\mu.$$

Tout comme les inégalités de Sobolev logarithmiques et de Poincaré (et leurs différentes variantes), les inégalités de transport permettent de démontrer des estimées de concentration.

Un de nos principaux résultats dans ce domaine est la caractérisation de l'inégalité T_1 complétant la caractérisation fonctionnelle établie par Bobkov-Götze. Nous montrons ensuite comment étendre aux suites dépendantes (notamment markoviennes) ces inégalités de transport et nous les appliquons au cas de lois de diffusions ainsi qu'à certains systèmes dynamiques aléatoires. Nous proposons ensuite une méthode de type GIRSANOV pour établir le transport T_2 dans le cas infini-dimensionnel de la mesure de WIENER ou de lois de diffusions, obtenant ainsi des résultats pour ce transport qui ne sont pas déduits d'une inégalité de Log- SOBOLEV.

Nous obtenons aussi la meilleure constante dans l'inégalité isopérimétrique de CHEEGER et certaines inégalités de transport-information et des inégalités de concentration plus précises pour la moyenne empirique

$$\frac{1}{t} \int_0^t g(X_s) ds,$$

où $(X_t : t \geq 0)$ est un processus de Markov réversible par rapport à une probabilité μ , de générateur infinitésimal \mathcal{L} . Ces inégalités sont des conséquences immédiates de l'estimation en norme Lipschitzienne de la solution de l'équation de Poisson.

3.2 Inégalités de coût du transport et applications

3.2.1 Principaux résultats sur T_p

Le résultat principal de cette section est la caractérisation de T_1 comme une intégrabilité gaussienne de la mesure μ .

Théorème 3.2.1 *Les propriétés suivantes sont équivalentes :*

1. $\mu \in T_1(C)$.
2. pour toute fonction lipschitzienne F μ -intégrable et en posant $\|F\|_{\text{Lip}} = \sup_{x \neq y} \frac{|F(x) - F(y)|}{d(x, y)}$, on a

$$(3.5) \quad \forall \zeta \in \mathbb{R}, \quad \int_E e^{\zeta(F - \int F d\mu)} d\mu \leq \exp(2\zeta^2 C \|F\|_{\text{Lip}}^2).$$

3. il existe $\delta > 0$ tel que $\iint e^{\delta d^2(x,y)} d\mu(x) d\mu(y) < +\infty$.

Plus précisément, en supposant l'inégalité $T_1(C)$, la dernière intégrale est finie pour tout $\alpha < C^{-1}$ et inversement en supposant cette intégrale finie, l'inégalité $T_1(C)$ est vraie pour

$$(3.6) \quad C \leq \inf_{\delta, x_0} \left(\frac{1}{\delta} \left(1 + \log \int e^{\delta d^2(x, x_0)} d\mu(x) \right) \right) < +\infty.$$

L'équivalence entre les points 1. et 2. provient de l'article de Bobkov-Götze [33]. Nous avons montré l'équivalence entre 2. et 3. dans [H-D5] en partant de l'idée simple suivante : étant données ξ et ξ' deux copies indépendantes de la loi μ et F une fonction lipschitzienne centrée par rapport à μ , on a alors

$$\mathbb{E} e^{\zeta F(\xi)} \leq \mathbb{E} e^{\zeta F(\xi)} \mathbb{E} e^{-\zeta F(\xi')} = \mathbb{E} e^{\zeta(F(\xi) - F(\xi'))}.$$

L'évaluation de la constante finale est celle obtenue par Bolley-Villani [36]. Notre méthode donne une constante légèrement plus élevée.

L'article [H-D5] propose également une étude de la stabilité des inégalités T_p par des transformations lipschitziennes et la tensorisation de ces inégalités pour des suites dépendantes.

3.2.2 Inégalité de transport pour des variables aléatoires dépendantes

Méthode de couplage

Soit \mathbb{P} une mesure de probabilité sur (E^n, \mathcal{B}^n) ($n \geq 2$). Pour tout $x \in E^n$, on note $x^i := (x_1, \dots, x_i)$. Soit $P_i(\cdot/x^{i-1})$ la loi conditionnelle régulière de x_i sachant x^{i-1} pour $i \geq 2$. Par convention $P_1(\cdot/x^0)$ est la loi de x_1 sous \mathbb{P} , où x^0 est un certain point fixé. Quand \mathbb{P} est Markov, alors $P_i(\cdot/x^{i-1}) = P_i(\cdot/x_{i-1})$ est le noyau de transition à l'étape $i - 1$.

Notre but ici est d'étendre les inégalités de coût de transport (3.2) pour des mesures de probabilités \mathbb{P} sur (E^n, d_{l_p}) , où

$$d_{l_p}(x, y) := \left(\sum_{i=1}^n d(x_i, y_i)^p \right)^{1/p}.$$

Théorème 3.2.2 Soit \mathbb{P} une mesure de probabilité sur E^n . Supposons que

(H1) il existe une constante $C > 0$ telle que pour tout $i \geq 1, x^{i-1}$ dans E^{i-1} et pour toute probabilité ν sur E , on ait

$$W_1^d(\nu, P_i(\cdot/x^{i-1})) \leq \sqrt{2CH(\nu/P_i(\cdot/x^{i-1}))},$$

(H2) il existe des constantes $a_j \geq 0$ avec $r := \sum_{j=1}^{\infty} a_j < 1$ telles que

$$W_1^d(P_i(\cdot/x^{i-1}), P_i(\cdot/\tilde{x}^{i-1})) \leq \sum_{j=1}^{i-1} a_j d(x_{i-j}, \tilde{x}_{i-j}), \text{ pour tout } i \geq 2, \text{ et tous } x^{i-1}, \tilde{x}^{i-1} \text{ dans } E^{i-1}.$$

Alors pour toute mesure de probabilité \mathbb{Q} sur E^n , on a

$$(3.7) \quad W_1^{d_{l_1}}(\mathbb{Q}, \mathbb{P}) \leq \frac{1}{1-r} \sqrt{2nCH(\mathbb{Q}/\mathbb{P})}.$$

Notons que pour toute fonction à valeurs réelles f sur E^n , $\|f\|_{\text{Lip}(d_{l_1})} \leq \alpha$ si pour tout $k = 1, \dots, n$

$$(3.8) \quad |f_k(x_k) - f_k(y_k)| \leq \alpha d(x_k, y_k), \quad \forall x_k, y_k \in E$$

où $f_k(x_k)$ est la fonction f par rapport à la k -ième variable alors que les autres sont constantes.

Alors grâce à l'article de Bobkov-Götze [33], on obtient, pour toute fonction réelle F sur E^n satisfaisant (3.8), l'inégalité :

$$\mathbb{E}_{\mathbb{P}} e^{\lambda(F - \mathbb{E}_{\mathbb{P}} F)} \leq \exp \left(\frac{\lambda^2 \alpha^2 n}{2(1-r)^2} \right), \quad \forall \lambda \in \mathbb{R}.$$

En particulier, on obtient pour tout $t > 0$:

$$(3.9) \quad \mathbb{P}(F > \mathbb{E}_{\mathbb{P}} F + t) \leq \exp \left(-\frac{t^2(1-r)^2}{2n\alpha^2} \right).$$

De manière similaire on a le théorème suivant :

Théorème 3.2.3 *Soit \mathbb{P} une mesure de probabilité sur E^n , et $1 \leq p \leq 2$.*

(H1') *il existe une constante $C > 0$ telle que pour tous $i \geq 1$, pour tous x^{i-1} dans E^{i-1} et pour toute probabilité ν sur E , on ait*

$$W_p^d(\nu, P_i(\cdot/x^{i-1})) \leq \sqrt{2CH(\nu/P_i(\cdot/x^{i-1}))}.$$

(H2') *il existe $a_j \geq 0$ avec $r^p := \sum_{j=1}^{\infty} (a_j)^p < 1$ telle que pour tout $i \geq 1$ et pour tous x^{i-1}, \tilde{x}^{i-1} dans E^{i-1}*

$$(3.10) \quad [W_p^d(P_i(\cdot/x^{i-1}), P_i(\cdot/\tilde{x}^{i-1}))]^p \leq \sum_{j=1}^{i-1} (a_j)^p d^p(x_{i-j}, \tilde{x}_{i-j}).$$

Alors pour toute mesure de probabilité \mathbb{Q} sur E^n ,

$$W_p^{d_{l_p}}(\mathbb{Q}, \mathbb{P}) \leq \left(\frac{a^{p-1}}{1 - r^p b^{p-1}} \right)^{1/p} \sqrt{2Cn^\beta H(\mathbb{Q}/\mathbb{P})},$$

où $\beta := \left(\frac{2}{p} - 1 \right)$ et les réels $a, b > 1$ forment une paire arbitraire de nombres conjugués ($1/a + 1/b = 1$) avec $r^p b^{p-1} < 1$. En particulier, quand $p = 2$, en optimisant suivant (a, b) , on obtient

$$W_2^{d_{l_2}}(\mathbb{Q}, \mathbb{P}) \leq \frac{1}{1-r} \sqrt{2CH(\mathbb{Q}/\mathbb{P})}.$$

La condition **(H1)** (ou **(H1')**) est le point de départ de notre tensorisation de $T_p(C)$. Elle est vérifiée par beaucoup d'exemples : l'EDS (3.17), les systèmes dynamiques aléatoires ou les champs de Gibbs. La condition **(H2)** (ou **(H2')**), qui signifie une dépendance très faible du présent par rapport au passé, est une condition cruciale.

La constante C_n pour l'inégalité de transport T_1 de \mathbb{P}_x croît linéairement avec la dimension n , ce qui est naturel, vu l'inégalité de Hoeffding (3.9). Ceci est complètement différent du cas de l'inégalité T_2 pour laquelle on espère que la constante reste indépendante de la dimension n , comme cela a déjà été vu dans le Théorème 3.2.3.

Sous la condition **(H1')** et l'inégalité (3.10) mais sans la condition de contraction $r^p := \sum_j (a_j)^p < 1$, on a toujours $\mathbb{P}_x \in T_p(C_n)$ sur E^n par rapport à la métrique d_{l_p} pour une certaine constante C_n (mais l'estimation cruciale de C_n dans le Théorème 3.2.3 est perdue).

Méthode martingale

L'inégalité (3.9) appliquée à $F(X_1, \dots, X_n) = \sum_{k=1}^n f(X_k)$ et à la métrique discrète d , dans le cas où les (X_k) sont indépendantes et $\|f(X_k)\|_\infty \leq \alpha$, devient exactement l'inégalité précise de Hoeffding. Mais quand on l'applique à $F(X_1, \dots, X_n) = f(X_n)$, elle ne donne pas le bon ordre de n pour n grand. Comme cette question est importante pour la connaissance de la mesure invariante, on utilise la méthode martingale de McDiarmid [128] pour obtenir le théorème :

Théorème 3.2.4 *Soit \mathbb{P} une mesure de probabilité sur E^n satisfaisant l'hypothèse **(H1)** du Théorème 3.2.2. Supposons au lieu de **(H2)** que*

(H3) *il existe une certaine constante $S > 0$ telle que pour toute fonction réelle bornée et Lipschitzienne $f(x_{k+1}, \dots, x_n)$ avec $\|f\|_{\text{Lip}(d_{l_1})} \leq 1$, pour tout $x \in E^n$ et tout $y_k \in E$, on a*

$$|\mathbb{E}_{\mathbb{P}}(f(X_{k+1}, \dots, X_n)/X^k = x^k) - \mathbb{E}_{\mathbb{P}}(f(X_{k+1}, \dots, X_n)/X^k = (x^{k-1}, y_k))| \leq Sd(x_k, y_k).$$

Alors, pour toute fonction F sur E^n satisfaisant (3.8), on a

$$(3.11) \quad \mathbb{E}_{\mathbb{P}} e^{\lambda(F - \mathbb{E}_{\mathbb{P}} F)} \leq \exp\left(\frac{C\lambda^2(1+S)^2\alpha^2 n}{2}\right), \quad \forall \lambda \in \mathbb{R}.$$

Où d'une manière équivalente $\mathbb{P} \in T_1(C_n)$ avec $C_n = n(1+S)^2$.

Si la condition **(H2)** est vue comme de type *backward*, alors **(H3)** peut être vue comme de type *forward*. En effet, **(H3)** est équivalente à

$$W_1^{d_{l_1}}(\mathbb{P}(dx_{k+1}^n/x_k, x^{k-1}), \mathbb{P}(dx_{k+1}^n/y_k, x^{k-1})) \leq Sd(x_k, y_k).$$

Cela signifie intuitivement que le présent n'a pas beaucoup d'influence sur l'avenir du processus \mathbb{P} . Dans des situations concrètes, **(H3)** est souvent plus faible que **(H2)**.

Quand $d(x, y) = 1_{x \neq y}$, **(H1)** est satisfaite, avec $C = 1/4$ et ce résultat est dû essentiellement à E. Rio [151]. En utilisant une condition différente de **(H3)**, E. Rio a aussi montré que la constante S de la condition **(H3)** vérifie $S \leq 2 \sum_{j=1}^{\infty} \phi_j$, où ϕ_j est le coefficient du mélange uniforme de la suite (X_n) .

3.2.3 Inégalité de transport pour des systèmes dynamiques aléatoires

Soit E une variété Riemannienne connectée et complète munie d'une métrique Riemannienne d . On considère maintenant le système dynamique linéaire aléatoire perturbé sur E ,

$$(3.12) \quad X_0(x) := x \in E, \quad X_{n+1} = F(X_n(x), W_{n+1}), \quad n \geq 0,$$

où le bruit $(W_n)_{n \geq 0}$ est une suite de v.a.r.i.i.d. à valeurs dans un certain espace mesurable (G, \mathcal{G}) , défini sur un certain espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, et où $F(x, w) : E \times G \rightarrow E$ est mesurable. On désigne par $P(x, dy)$ la loi de $F(x, W_1)$.

Proposition 3.2.5 *Supposons qu'il existe $\delta > 0$ tel que*

$$(3.13) \quad \sup_{x \in E} \mathbb{E} \left(e^{\delta d(F(x, W_1), F(x, W_2))^2} \right) < +\infty.$$

On suppose qu'il existe $0 \leq r < 1$ tel que

$$(3.14) \quad \mathbb{E}(d(F(x, W_1), F(\tilde{x}, W_1))) \leq r d(x, \tilde{x}), \quad \forall x, \tilde{x} \in E,$$

ou bien, plus généralement, qu'il existe une constante $S \geq 0$, telle que :

$$(3.15) \quad \sum_{n=1}^{\infty} \mathbb{E}(d(X_n(x), X_n(\tilde{x}))) \leq S d(x, \tilde{x}), \quad \forall x, \tilde{x} \in E.$$

Alors il existe une constante $C > 0$ telle que, pour tout $n \geq 1$ et pour toute probabilité \mathbb{Q}^n sur E^n , on ait

$$W_1^{d_{l_1}}(\mathbb{Q}^n, \mathbb{P}_x^n) \leq \sqrt{CnH(\mathbb{Q}^n/\mathbb{P}_x^n)},$$

où \mathbb{P}_x^n est la loi de $(X_k(x))_{1 \leq k \leq n}$ sur E^n .

Pour comparer (3.15) avec (3.14), on considère le modèle linéaire $F(x, w) = Ax + w$ où $E = \mathbb{R}^d = G$, $A \in \mathcal{M}_{d \times d}$. Dans ce cas, la condition (3.14) veut dire que $r = \|A\| := \sup\{|Ax|; |x| \leq 1\} < 1$, alors que (3.15) devient équivalent à $\max\{|\lambda|; \lambda \text{ rayon spectral } A\} < 1$ ce qui est beaucoup plus faible. Cette dernière condition est bien connue pour être optimale pour l'ergodicité de ce modèle linéaire ARMA.

Considérons le cas particulier du modèle précédent où $F(x, w) = f(x) + \sigma(x)w$, $E = \mathbb{R}^d$, $G = \mathbb{R}^n$, $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times n}$,

$$X_0(x) = x, \quad X_{n+1}(x) = f(X_n(x)) + \sigma(X_n(x))W_{n+1},$$

et où le bruit $(W_n)_{n \in \mathbb{Z}}$ est tel que $\mathbb{E}W_1 = 0$. Supposons

- (i) $\mathbb{P}_W := \mathbb{P}(W_1 \in \cdot) \in T_2(C)$ sur \mathbb{R}^n par rapport à la métrique euclidienne ;
- (ii) $|\sigma(x)w| \leq K|w|$, $\forall (x, w) \in \mathbb{R}^d \times \mathbb{R}^n$;
- (iii) pour un certain $r \in [0, 1)$, on a

$$(3.16) \quad \sqrt{|f(x) - f(\tilde{x})|^2 + \mathbb{E}|(\sigma(x) - \sigma(\tilde{x}))W_1|^2} \leq r|x - \tilde{x}|, \quad \forall x, \tilde{x} \in \mathbb{R}^d.$$

Notons que les conditions (i) et (ii) impliquent que $P(x, \cdot) \in T_2(CK^2)$ pour tout $x \in \mathbb{R}^d$ et que la condition (iii) implique **(H2')** avec le même r , pour $p = 2$. Donc, grâce au Théorème 3.2.4, $\mathbb{P}_x^n \in T_2(CK^2/(1-r)^2)$. Grâce aux résultats de Bobkov-Gentil-Ledoux [32], cela donne, pour toute fonction mesurable $F(x_1, \dots, x_n) \in L^1((\mathbb{R}^d)^n, \mathbb{P}_x^n)$, l'inégalité

$$\mathbb{E} \exp(\rho QF(X_1(x), \dots, X_n(x))) \leq \exp(\rho \mathbb{E} F(X_1(x), \dots, X_n(x)))$$

où

$$\rho := \frac{(1-r)^2}{CK^2}, \quad QF(x_1, \dots, x_n) := \inf_{y \in (\mathbb{R}^d)^n} \left(F(x+y) + \frac{1}{2} \sum_{k=1}^n |y_k|^2 \right).$$

Comme il apparaît déjà dans [32], plusieurs estimations des intégrales de Laplace sont des conséquences de versions des inégalités fonctionnelles de type $T_2(C)$. Pour l'instant, le corollaire 6.1 dans [32] dit que pour toute fonction convexe F sur $(\mathbb{R}^d)^n$, on a

$$\mathbb{E}_{\mathbb{P}_x^n} \exp \left(\rho \left[F - \frac{1}{2} \sum_{k=1}^n (\partial_k F)^2 \right] \right) \leq \exp(\rho \mathbb{E}_{\mathbb{P}_x^n} F).$$

Inégalité T_1 pour une EDS

Considérons (X_t) défini par l'EDS

$$(3.17) \quad dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x,$$

où σ et b sont globalement lipschitziennes et vérifient :

$$(3.18) \quad \sup_x \|\sigma(x)\| \leq A, \quad \langle y - x, b(y) - b(x) \rangle \leq B(1 + |y - x|^2),$$

et

$$(3.19) \quad 2\langle y - x, b(y) - b(x) \rangle + \text{tr}((\sigma(y) - \sigma(x))(\sigma(y) - \sigma(x))^*) \leq -\delta|y - x|^2, \quad \text{avec } \delta > 0.$$

Alors la loi de $(X_t(x))_{t \in [0, n]}$ satisfait une inégalité $T_1(Cn)$ par rapport à la métrique définie par :

$$d(\gamma_1, \gamma_2) = \sum_{k=0}^{n-1} \sup_{k \leq t \leq k+1} |\gamma_1(t) - \gamma_2(t)|.$$

Soit (P_t) le semi-groupe de transition associé à la diffusion (X_t) . Notons que sous la condition (3.19), on a

$$\mathbb{E}|X_t(x) - X_t(\tilde{x})|^2 \leq |x - \tilde{x}|^2 e^{-2\delta t}, \quad \forall t \geq 0.$$

Ceci entraîne non seulement l'existence et l'unicité de la mesure de probabilité invariante μ de (P_t) , mais aussi que $P_t(x, \cdot) \rightarrow \mu$ quand $t \rightarrow \infty$ pour tout $x \in \mathbb{R}^d$.

Présentons l'inégalité de type Hoeffding pour $F(\gamma) := \int_0^n V(\gamma(t))dt$ où $V : \mathbb{R}^d \rightarrow \mathbb{R}$ satisfait $\|V\|_{\text{Lip}} \leq \alpha$. Pour un tel V , $\|F\|_{\text{Lip}} \leq \alpha$ par rapport à la métrique précédente. Le travail de Bobkov-Götze [33] entraîne donc :

$$\mathbb{P} \left(\int_0^n [V(X_t(x)) - \mathbb{E}V(X_t(x))]dt > r \right) \leq \exp \left(-\frac{r^2}{2nC} \right), \quad \forall r > 0.$$

Le seul désavantage de l'inégalité précédente est que la constante C n'est pas bien déterminée.

Inégalité T_2 pour une EDS

On sait que la mesure Gaussienne satisfait T_2 , qu'une inégalité de Sobolev logarithmique implique T_2 , et que l'implication inverse est vraie si la mesure considérée est une mesure de Boltzmann issue d'un potentiel convexe ϕ (plus une condition sur la constante du transport et la minoration de la Hessienne de ϕ). On sait que T_2 implique à son tour T_1 et une inégalité de Poincaré. Dans ce paragraphe, notre point de vue est le suivant : que peut-on dire sur T_2 sans inégalité de Sobolev logarithmique ?

Si nous reprenons l'exemple de la diffusion précédente avec des coefficients localement lipschitziens satisfaisant (3.19) et tel que $\|\sigma\|_\infty$ soit finie, alors la loi de $X_t(x)$ satisfait une inégalité $T_2(\|\sigma\|_\infty^2/\delta^2)$ pour tout point de départ x et à tout temps T par rapport à la métrique L^2 donnée par

$$(3.20) \quad d_2(\gamma_1, \gamma_2) = \sqrt{\int_0^T |\gamma_1(t) - \gamma_2(t)|^2 dt}.$$

En outre, on a $P_T(x, \cdot) \in T_2\left(\frac{\|\sigma\|_\infty^2}{2\delta}\right)$ sur \mathbb{R}^d , ainsi que l'unique mesure de probabilité invariante μ de (P_t) .

Pour la loi de cette diffusion, aucune inégalité de Sobolev logarithmique n'est connue. En fait T_2 n'a été considérée pour des mesures de processus que pour la mesure de Wiener et la distance de Cameron-Martin, ou pour des lois d'e.d.s. qui sont des transformations Lipschitziennes de la mesure de Wiener, (voir cependant Wang [168] pour une extension aux inégalités fini-dimensionnelles). La preuve repose ici sur une transformation de Girsanov permettant de construire un couplage pour la distance de Wasserstein, puis de revenir par un contrôle direct sur l'entropie en évaluant la distance L^2 de ce couplage. Cet exemple (donné dans [H-D5]) n'est cependant pas satisfaisant dans le sens où l'inégalité de Sobolev logarithmique n'est pas connue et ne répond donc pas à la question principale : y a-t-il une différence entre T_2 et log-Sobolev ?

Revenons à l'e.d.s. (3.17). Supposons que $\|\nabla b\| \leq K$, $n = d$ et $\sigma(x) = \sigma = I_d$. Capitaine-Hsu-Ledoux, dans [42], montrent l'inégalité de log-Sobolev de la loi \mathbb{P}_x de la diffusion sur l'espace de Cameron-Martin associé au gradient de Malliavin. Comme le suggère le résultat de Otto-Villani [141], l'inégalité de log-Sobolev implique l'inégalité $T_2(C)$ et on doit avoir $\mathbb{P}_x \in T_2(C)$ sur $C([0, T])$ par rapport à la pseudo métrique

$$(3.21) \quad d_H(\gamma_1, \gamma_2) := \begin{cases} \|\gamma_1 - \gamma_2\|_H, & \text{si } \gamma_1 - \gamma_2 \in H; \\ +\infty & \text{sinon} \end{cases}$$

Nous avons étendu l'inégalité T_2 due à Talagrand pour les mesures Gaussiennes, à la mesure de Wiener \mathbb{P} sur $C([0, T], \mathbb{R}^d)$. En utilisant la formule de Girsanov, on montre que $\mathbb{P} \in T_2(1)$ sur $(C([0, T], \mathbb{R}^d), d_H)$. La constante est optimale. L'inégalité de transport T_2 de la mesure de Wiener par rapport à la métrique de Cameron-Martin est mentionnée dans la thèse de Gentil [83]. Sa preuve n'utilise pas notre formalisme.

Avec les mêmes arguments, on obtient une inégalité T_2 pour la solution de l'e.d.s. (3.17) avec $\sigma = I_d$, par rapport à la métrique de Cameron-Martin d_H . Plus précisément, en supposant que $\nabla^s b := (\frac{1}{2}(\partial_i b^j + \partial_j b^i))_{1 \leq i, j \leq d} \leq BI_d$ et $\|\nabla b\| \leq K$, alors on obtient pour tout point initial x , $\mathbb{P}_x \in T_2(\alpha^2)$ sur $C([0, T], \mathbb{R}^d)$ par rapport à la métrique d_H , où α^2 est donné par

$$\alpha^2 := \alpha^2(T, K, B) = \begin{cases} 2\left(1 + \frac{K^2}{B^2}\right), & \text{si } B < 0 \\ 2\left(1 + K^2 \frac{e^{2BT}}{2B^2}\right), & \text{si } B > 0 \\ 2\left(1 + \frac{K^2 T^2}{2}\right), & \text{si } B = 0. \end{cases}$$

Ce résultat entraîne une bonne estimation pour des fonctionnelles comme $F(\gamma) = \sup_{t \leq T} |X_t(x) - x|$ pour T petit (estimations de type Varadhan-Freidlin-Ventzell). Cependant comme Wu l'a déjà noté dans

[174], ce type d'inégalités (même avec α^2 indépendant de T), formellement plus faible que l'inégalité de log-Sobolev correspondante, n'entraîne pas une estimation d'ordre optimal pour des fonctionnelles additives comme $\int_0^T V(X_t)dt$ pour T grand. Pour obtenir une meilleure estimation, on utilise l'inégalité T_2 par rapport à la métrique d_2 .

Il est intéressant de voir si nos résultats restent valables dans le cas où $\nabla^s b \leq -BI_d$ sans la condition $\|\nabla b\| \leq K$.

Comme indiqué dans [32], des conséquences intéressantes peuvent être déduites de ce résultat comme par exemple le corollaire suivant :

Corollaire 3.2.6 *On a pour tout $T > 0$,*

(a) *pour toute fonction cylindrique régulière F sur $G := L^2([0, T], dt; \mathbb{R}^d) \supset C([0, T], \mathbb{R}^d)$, c.à.d.*

$$F \in \mathcal{FC}_b^\infty := \{f(\langle \gamma, h_1 \rangle, \dots, \langle \gamma, h_n \rangle); n \geq 1, h_i \in \tilde{H}, f \in C_b^\infty(\mathbb{R}^n)\}$$

(où $\langle \gamma_1, \gamma_2 \rangle := \int_0^T \gamma_1(t)\gamma_2(t)dt$), l'inégalité de Poincaré suivante a lieu :

$$(3.22) \quad \text{Var}_{\mathbb{P}_x}(F) \leq \frac{\|\sigma\|_\infty^2}{\delta^2} \int_{C([0, T], \mathbb{R}^d)} \|\nabla F(\gamma)\|_G^2 d\mathbb{P}_x(\gamma)$$

où $\text{Var}_{\mathbb{P}_x}(F)$ est la variance de F sous la loi \mathbb{P}_x , et $\nabla F(\gamma) \in G$ est le gradient de F en γ .

(b) *pour toute fonction $g \in C_b^\infty(\mathbb{R}^d)$,*

$$(3.23) \quad \text{Var}_{P_T(x, \cdot)}(g) \leq \frac{\|\sigma\|_\infty^2}{2\delta} \int_{\mathbb{R}^d} |\nabla g(y)|^2 P_T(x, dy).$$

(c) (inégalité de type Tsirel'son) *pour tout sous-ensemble non-vide K de G tel que $Z(\gamma) := \sup_{h \in K} \langle \gamma, h \rangle \in L^1(\mathbb{P}_x)$, alors*

$$(3.24) \quad \int \exp\left(\frac{\delta^2}{\|\sigma\|_\infty^2} \sup_{h \in K} [\langle \gamma, h \rangle - |h|_G^2/2]\right) d\mathbb{P}_x \leq \exp\left(\frac{\delta^2}{\|\sigma\|_\infty^2} \mathbb{E}^{\mathbb{P}_x} Z\right).$$

(d) (inégalité de type Hoeffding) *pour toute fonction $V : \mathbb{R}^d \rightarrow \mathbb{R}$ telle que $\|V\|_{Lip} \leq \alpha$,*

$$\mathbb{P}\left(\frac{1}{T} \int_0^T V(X_t(x))dt - \mathbb{E} \frac{1}{T} \int_0^T V(X_t(x))dt > r\right) \leq \exp\left(-\frac{Tr^2 \|\sigma\|_\infty^2}{2\alpha^2 \delta^2}\right), \quad \forall r > 0.$$

3.3 Inégalités fonctionnelles par l'estimation en norme Lipschitzienne d'une équation de Poisson

Soit I un intervalle de \mathbb{R} . Son intérieur est donc de la forme $I^0 = (x_0, y_0)$ où $-\infty \leq x_0 < y_0 \leq +\infty$. On considère l'opérateur de Sturm-Liouville sur I :

$$\mathcal{L} = a(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx}$$

avec des conditions de type Neumann sur la frontière $\partial I = \{x_0, y_0\} \cap \mathbb{R}$, où

(A1) les fonctions $a, b : I \rightarrow \mathbb{R}$ sont mesurables et localement bornées $a(x) > 0$ dx -p.s. et $1/a$ est localement dx -intégrable sur I , (ici, dx est la mesure de Lebesgue).

Sur I^0 , \mathcal{L} peut se réécrire sous la forme de Feller suivante :

$$\mathcal{L} = \frac{1}{m'(x)} \frac{d}{dx} \left(\frac{1}{s'(x)} \frac{d}{dx} \right) = \frac{d}{dm} \frac{d}{ds}$$

où m et s sont respectivement la mesure de vitesse et la fonction d'échelle de Feller, qui sont des fonctions absolument continues sur I telles que pour un certain point c fixé dans I on ait, $dx - p.s.$

$$s'(x) = \exp\left(-\int_c^x \frac{b(u)}{a(u)} du\right) \quad \text{et} \quad m'(x) = \frac{1}{a(x)s'(x)}$$

Soit $C_0^\infty(I)$ l'espace des fonctions réelles f indéfiniment différentiables sur I à support compact et $\mathcal{D} = \{f \in C_0^\infty(I); f'|_{\partial I} = 0\}$.

L'opérateur \mathcal{L} , défini sur \mathcal{D} , est symétrique sur $L^2(I, m)$ (m indique aussi la mesure $m'(x)dx$).

Soit $(X_t : t \geq 0)$ la diffusion sur l'intervalle I de générateur infinitésimal \mathcal{L} . Les conditions de Neumann sur le bord correspondent à une réflexion au bord ∂I . Voir [104] pour plus de détails et pour des définitions précises. On suppose que :

(A2) la diffusion est non explosive et positivement récurrente, c.à.d., $m(I) = \int_I m'(y)dy < +\infty$ et

$$\int_c^{y_0} s'(x) \left(\int_c^x m'(y)dy \right) dx = +\infty \quad \text{si } y_0 \notin I \quad \text{et} \quad \int_{x_0}^c s'(x) \left(\int_{x_0}^x m'(y)dy \right) dx = +\infty \quad \text{si } x_0 \notin I$$

(A3) le générateur \mathcal{L} , défini sur \mathcal{D} , est essentiellement auto-adjoint sur $L^2(I, dm)$, où d'une manière équivalente (voir [T-D18], [74]) :

$$s \notin L^2((x_0, c], dm), \quad \text{si } x_0 \notin I; \quad \text{et } s \notin L^2([c, y_0), dm), \quad \text{si } y_0 \notin I.$$

Notons que si $a(x) = 1$ et $I = \mathbb{R}$, les hypothèses **(A2)** et **(A3)** sont automatiquement satisfaites dès que $m(I) < +\infty$ (voir [104] pour **(A2)**, [T-D18] et [74] pour **(A3)**).

Dans la suite, on suppose que les conditions **(A1)**–**(A3)** sont satisfaites.

Dans ce cas $(X_t)_{t \geq 0}$ est réversible par rapport à la mesure de probabilité $\mu(dx) = \frac{1}{m(I)} m'(x)dx$. Soient $(P_t)_{t \geq 0}$ le semi-groupe de transition de $(X_t)_{t \geq 0}$, \mathcal{L}_p le générateur de (P_t) dans $L^p(I, \mu)$ de domaine $\mathbb{D}(\mathcal{L}_p)$ (pour $p \in [1, +\infty)$), ce générateur est une extension de $(\mathcal{L}, \mathcal{D})$.

Pour $g \in L^2(I, \mu)$ telle que $\mu(g) = 0$, on considère l'équation de Poisson

$$(3.25) \quad -\mathcal{L}_2 G = g.$$

Par l'ergodicité de la diffusion, la solution G de l'équation de Poisson, si elle existe, est unique dans $L^2(I, \mu)$ à une constante près.

Dans l'interprétation physique de la diffusion de la chaleur, g représente la source de la chaleur, G est la distribution de la chaleur d'équilibre. L'objectif de ce travail est d'estimer

$$(3.26) \quad \|G\|_{Lip(\rho)} := \sup_{x, y \in I, x < y} \frac{|G(y) - G(x)|}{|\rho(y) - \rho(x)|}$$

en fonction des différentes normes de la source de chaleur g . Ici ρ est une certaine fonction croissante et absolument continue sur I telle que $\rho'(x) > 0$ $dx - p.s.$

Soit λ_1 le trou spectral de \mathcal{L}_2 , c.à.d. la plus petite valeur propre ou le point spectral au-dessus de zéro de $-\mathcal{L}_2$. Alors $c_P := \lambda_1^{-1}$ est la meilleure constante dans l'inégalité de Poincaré suivante

$$(3.27) \quad \text{Var}_\mu(f) \leq c_P \int_I a(x) f'(x)^2 d\mu(x), \quad \forall f \in \mathcal{D}.$$

L'importance du trou spectral réside dans le fait qu'il décrit le taux de la convergence exponentielle

$$\|P_t f - \mu(f)\|_2 \leq e^{-\lambda_1 t} \|f - \mu(f)\|_2, \quad \forall t \geq 0$$

où $\|\cdot\|_2$ est la norme dans $L^2(I, \mu)$. La constante λ_1 peut aussi être interprétée *via* l'équation de Poisson de la manière suivante :

$$\|G - \mu(G)\|_2 \leq c_P \|g\|_2 \quad \text{ou} \quad \int_I a(x) G'(x)^2 d\mu(x) \leq c_P \|g\|_2^2.$$

Les interprétations physiques expliquent l'importance fondamentale de l'étude de λ_1 ou de c_P . Comme l'étude de λ_1 a une très longue histoire, il n'est pas possible pour nous de la décrire même dans ses grandes lignes. Le lecteur intéressé consultera les ouvrages [44, 168] pour des bibliographies complètes. Pour l'inégalité plus forte de log-Sobolev, la première caractérisation est due à Bobkov-Götze [29] (voir aussi [77, 44] pour de nouvelles améliorations de la constante).

Notre premier objectif est de comprendre la formule variationnelle due à Chen pour λ_1 ([43]) :

$$(3.28) \quad c_P = \inf_{\rho} \sup_{x \in I} \frac{s'(x)}{\rho'(x)} \int_x^{y_0} [\rho(t) - \mu(\rho)] m'(t) dt,$$

où ρ parcourt toutes les fonctions de $C^1(I)$ avec $\rho' > 0$.

Remarquons que l'on ne connaît aucune formule variationnelle pour la meilleure constante dans l'inégalité de log-Sobolev sur la droite réelle.

Notre motivation principale vient en fait de quelques inégalités de concentration pour la moyenne empirique

$$(1/t) \int_0^t g(X_s) ds,$$

qui sont des conséquences immédiates de l'estimation de $\|G\|_{Lip(\rho)}$ *via* la décomposition forward-backward de martingales ou les inégalités d'information-transport développées dans [95].

Notre méthode d'estimation de $\|G\|_{Lip(\rho)}$ est directe. La solution de l'équation de Poisson peut être calculée explicitement (contrairement à celle de l'équation de la chaleur correspondante) et quelques contrôles supplémentaires (faciles) sont nécessaires pour finir le travail.

En plus de ces motivations, l'évaluation de G' est physiquement significative : dans le problème de diffusion de la chaleur, en présence de la source de la chaleur g avec $\mu(g) = 0$, G représente la distribution de la chaleur d'équilibre. Une estimation sur $|G'|$ nous permet donc de contrôler la variation de la distribution de la chaleur d'équilibre.

3.3.1 Principaux résultats

Étant donné une fonction absolument continue $\rho : I \rightarrow \mathbb{R}$ telle que $\rho' > 0, dx - p.s.$, on considère $d_\rho(x, y) = |\rho(x) - \rho(y)|$ la métrique sur I associée à ρ .

Si la norme Lipschitzienne $\|f\|_{Lip(\rho)}$ de f par rapport à d_ρ définie dans (3.26) est finie, on dit que f est ρ -Lipschitzienne.

Soit $L_0^2(I, \mu) := \{f \in L^2(I, \mu); \mu(f) = 0\}$. Nous pouvons maintenant énoncer notre premier résultat.

Théorème 3.3.1 *Soient ρ, ρ_1, ρ_2 des fonctions absolument continues sur I telles que $\rho, \rho_k \in L^2(I, \mu)$, $\rho', \rho'_k > 0, dx - p.s.$*

1. *Si*

$$(3.29) \quad c_{Lip}(\rho_1, \rho_2) := \operatorname{ess\,sup}_{x \in I} \frac{s'(x)}{\rho_2'(x)} \int_x^{y_0} [\rho_1(t) - \mu(\rho_1)] m'(t) dt < +\infty,$$

alors pour toute fonction $g \in L_0^2(I, \mu)$ qui est ρ_1 -Lipschitzienne, il existe une unique solution G avec $\mu(G) = 0$ appartenant au domaine $\mathbb{D}(\mathcal{L}_2)$ de l'équation de Poisson (3.25).

De plus G est ρ_2 -Lipschitzienne et satisfait

$$(3.30) \quad \|G\|_{Lip(\rho_2)} \leq c_{Lip}(\rho_1, \rho_2) \|g\|_{Lip(\rho_1)}.$$

En outre cette inégalité (3.30) devient une égalité pour $g = \rho_1 - \mu(\rho_1)$.

2. Soit $\varphi : I \rightarrow \mathbb{R}^+$ une fonction positive sur $L^2(I, \mu)$. Si

$$(3.31) \quad c(\varphi, \rho) := \operatorname{ess\,sup}_{x \in I} \frac{s'(x)}{\rho'(x)} m(I) \left(\mu(I_x^+) \int_{I_x^-} \varphi d\mu + \mu(I_x^-) \int_{I_x^+} \varphi d\mu \right) < +\infty,$$

où $I_x^+ = \{y \in I; y \geq x\}$, $I_x^- = \{y \in I; y < x\}$, alors pour toute fonction $g \in L^2(I, \mu)$ telle que $|g| \leq \varphi$, il existe une unique solution G avec $\mu(G) = 0$ de l'équation Poisson $-\mathcal{L}_2 G = g - \mu(g)$. En outre G est ρ -Lipschitzienne et satisfait

$$(3.32) \quad \sup_{g: |g| \leq \varphi} \|G\|_{\text{Lip}(\rho)} = c(\varphi, \rho).$$

Soit $C_{\text{Lip}(\rho),0}$ l'espace de Banach des fonctions g avec $\mu(g) = 0$ qui sont ρ -Lipschitziennes, muni de la norme $\|\cdot\|_{\text{Lip}(\rho)}$. La partie 1. du théorème précédent dit que l'opérateur de Poisson $(-\mathcal{L}_2)^{-1} : C_{\text{Lip}(\rho_1),0} \rightarrow C_{\text{Lip}(\rho_2),0}$ est borné et que :

$$(3.33) \quad \|(-\mathcal{L}_2)^{-1}\|_{C_{\text{Lip}(\rho_1),0} \rightarrow C_{\text{Lip}(\rho_2),0}} = c_{\text{Lip}}(\rho_1, \rho_2).$$

Comme \mathcal{L}_2 est auto-adjoint sur $L_0^2(I, \mu)$, un résultat général d'analyse fonctionnelle (voir [172]) fournit :

$$\|(-\mathcal{L}_2)^{-1}\|_{L_0^2(I, \mu)} \leq \|(-\mathcal{L}_2)^{-1}\|_{C_{\text{Lip}(\rho),0} \rightarrow C_{\text{Lip}(\rho),0}},$$

dont le membre de gauche est exactement la constante de Poincaré c_P . On obtient donc :

$$c_P \leq \|(-\mathcal{L}_2)^{-1}\|_{C_{\text{Lip}(\rho),0} \rightarrow C_{\text{Lip}(\rho),0}} = c_{\text{Lip}}(\rho, \rho)$$

qui est exactement la majoration de c_P dans (3.28). C'est notre interprétation de la formule variationnelle de Chen (3.28).

Soit $\|g\|_\varphi$ la plus petite constante c telle que $|g(x)| \leq c\varphi(x)$ sur I et $b_\varphi \mathcal{B}$ l'espace de Banach de telles fonctions mesurables g de norme $\|g\|_\varphi$ finie. Soit $Pg = g - \mu(g) : L^2(I, \mu) \rightarrow L_0^2(I, \mu)$, la projection orthogonale. La partie 2. du théorème précédent entraîne que $(-\mathcal{L})^{-1}P$ est borné de $b_\varphi \mathcal{B}$ dans $C_{\text{Lip}(\rho),0}$ et sa norme est exactement $c(\varphi, \rho)$.

L'inégalité L^1 de Poincaré et l'inégalité isopérimétrique de Cheeger

L'inégalité de Poincaré a une contrepartie dans L^1 , en lien avec l'inégalité isopérimétrique de Cheeger. Soit $c_{P,1}$ la meilleure constante telle que l'inégalité L^1 de Poincaré ait lieu : pour toute fonction $f \in \mathcal{AC}(I) \cap L^1(I, \mu)$,

$$(3.34) \quad \int_I |f - \mu(f)| d\mu \leq c_{P,1} \int_I \sqrt{a(x)} |f'| d\mu$$

où $\mathcal{AC}(I)$ est l'espace de toutes les fonctions absolument continues sur I .

Soit c_{cheeger} la constante isopérimétrique de Cheeger de μ par rapport à la métrique intrinsèque d_{ρ_0} associée à la diffusion X_t avec $\rho_0(x) = \int_c^x \frac{1}{\sqrt{a(y)}} dy$, c.à.d. la meilleure constante dans l'inégalité isopérimétrique de Cheeger suivante :

$$\min(\mu(A), 1 - \mu(A)) \leq c_{\text{cheeger}} \mu_s(\partial A),$$

pour tous les sous-ensembles $A \subset I$. Ici ∂A est le bord de A et la mesure de surface μ_s de A est définie par $\mu_s(\partial A) = \liminf_{\varepsilon \rightarrow 0+} \frac{\mu(A_\varepsilon) - \mu(A)}{\varepsilon}$ avec $A_\varepsilon = \{x \in I, d_{\rho_0}(x, A) \leq \varepsilon\}$ le ε -voisinage de A .

Il est bien connu (cf. [30, 115]) que c_{cheeger} est aussi la meilleure constante dans la version fonctionnelle de l'inégalité isopérimétrique de Cheeger suivante : pour toute fonction $f \in \mathcal{AC}(I) \cap L^1(I, \mu)$

$$(3.35) \quad \int_I |f - m_\mu(f)| d\mu \leq c_{\text{cheeger}} \int_I \sqrt{a(x)} |f'| d\mu,$$

où $m_\mu(f)$ est une médiane de f par rapport à μ (via la formule de la co-aire). De l'encadrement suivant

$$\frac{1}{2}\mu(|f - \mu(f)|) \leq \mu(|f - m_\mu(f)|) \leq \mu(|f - \mu(f)|),$$

on tire

$$(3.36) \quad \frac{1}{2}c_{P,1} \leq c_{cheeger} \leq c_{P,1}.$$

Les deux inégalités précédentes sont toutes les deux optimales (*cf.* les exemples suivants).

Un résultat important de Bobkov-Houdré [30] dit que

$$(3.37) \quad c_{cheeger} = \operatorname{ess\,sup}_{x \in I} \frac{m(I) \min\{\mu(I_x^+), \mu(I_x^-)\}}{m'(x) \sqrt{a(x)}} = \operatorname{ess\,sup}_{x \in I} m(I) \sqrt{a(x)} s'(x) \min\{\mu(I_x^+), \mu(I_x^-)\},$$

ce qui veut dire grossièrement que l'ensemble extrême pour $c_{cheeger}$ est un demi intervalle I_x^+ .

Ces dernières années, la meilleure constante $c_{cheeger}$ (dans le cas multi-dimensionnel) a été étudiée de manière intensive (voir [41], [30], [115], [130], [179] pour des références pertinentes).

Le théorème 3.3.1 nous permet d'identifier la meilleure constante $c_{P,1}$ dans l'inégalité L^1 de Poincaré donnée dans (3.34). C'est le propos du théorème suivant

Théorème 3.3.2 *La meilleure constante $c_{P,1}$ dans l'inégalité L^1 de Poincaré (3.34) est finie ssi*

$$(3.38) \quad c_\delta := 2 \operatorname{ess\,sup}_{x \in I} \sqrt{a(x)} s'(x) m(I) \mu(I_x^+) \mu(I_x^-) < +\infty,$$

où I_x^+, I_x^- sont donnés dans le Théorème 3.3.1. Dans ce cas $c_{P,1} = c_\delta$.

En appliquant (3.34) à $(1 - nd_{\rho_0}(x, A))^+$ et en faisant tendre $n \rightarrow \infty$, on obtient l'inégalité isopérimétrique suivante

$$2\mu(A)\mu(A^c) \leq c_{P,1}\mu_s(\partial A)$$

pour tout sous-ensemble mesurable A de I .

Mais on ne sait pas si l'inégalité isopérimétrique précédente implique (3.34) dans le cas multidimensionnel (l'approche *via* la formule de la co-aire ne marche pas.) Cependant, la meilleure constante dans la dernière inégalité isopérimétrique est bornée supérieurement par $c_{P,1}$ et inférieurement par c_δ (juste en prenant $A = I_x^+$), et donc est égale à $c_\delta = c_{P,1}$ (par le Théorème 3.3.2) dans le cas uni-dimensionnel.

3.3.2 Applications aux inégalités de transport et de concentration

La forme de Dirichlet $(\mathcal{E}, (\mathcal{E}))$ associée au semi-groupe de transition (P_t) de (X_t) est donnée par :

$$\begin{aligned} \mathbb{D}(\mathcal{E}) &= \mathbb{D}(\sqrt{-\mathcal{L}_2}) = \{f \in L^2(I, \mu) \cap \mathcal{AC}(I), \int_I a(x) f'(x)^2 d\mu(x) < +\infty\}, \\ \mathcal{E}(f, f) &:= \int_I a(x) f'(x)^2 d\mu(x), \quad f \in \mathbb{D}(\mathcal{E}). \end{aligned}$$

Pour $f, g \in \mathbb{D}(\mathcal{E})$, soit $\Gamma(f, g) = af'g'$ l'opérateur carré-du-champ. L'information de Fisher-Donsker-Varadhan de ν par rapport à μ est définie par :

$$(3.39) \quad I(\nu|\mu) = \begin{cases} \mathcal{E}\left(\sqrt{\frac{d\nu}{d\mu}}, \sqrt{\frac{d\nu}{d\mu}}\right), & \text{si } \nu \ll \mu \text{ et } \sqrt{\frac{d\nu}{d\mu}} \in \mathbb{D}(\mathcal{E}); \\ +\infty, & \text{sinon.} \end{cases}$$

Rappelons que pour $\rho_0(x) = \int_c^x \frac{1}{\sqrt{a(y)}} dy$, la métrique associée $d_{\rho_0}(x, y) = |\rho_0(y) - \rho_0(x)|$ est la métrique intrinsèque de la diffusion (X_t) .

Corollaire 3.3.3 Soit $\rho \in \mathcal{AC}(I) \cap L^2(I, \mu)$ telle que $\rho'(x) > 0$, $dx - p.s.$ et soit

$$(3.40) \quad c_\rho = \operatorname{ess\,sup}_{x \in I} s'(x) \sqrt{a(x)} \int_x^{y_0} [\rho(y) - \mu(\rho)] m'(y) dy < +\infty.$$

Alors pour toute mesure $\nu \in \mathcal{M}_1(I)$, on a

$$(3.41) \quad (W_{1,d_\rho}(\nu, \mu))^2 \leq 4c_\rho^2 I(\nu|\mu),$$

ou d'une manière équivalente, pour toute fonction ρ -Lipschitzienne g sur I , on a, pour toute mesure initiale $\nu \ll \mu$ et tout réels $t, r > 0$,

$$(3.42) \quad \mathbb{P}_\nu \left(\frac{1}{t} \int_0^t g(X_s) ds > \mu(g) + r \right) \leq \left\| \frac{d\nu}{d\mu} \right\|_{L^2(I, \mu)} \exp \left(- \frac{tr^2}{4c_\rho^2 \|g\|_{Lip(\rho)}^2} \right).$$

La seconde inégalité peut être réécrite de la manière suivante :

$$W_{1,\rho}(f\mu, \mu) \leq c_\rho \int_I \sqrt{\Gamma(f, f)} d\mu \leq 2c_\rho \sqrt{I(f\mu|\mu)}.$$

En répétant les arguments précédents, mais en utilisant la partie 2. du Théorème 3.3.1, on obtient le corollaire ci-dessous.

Corollaire 3.3.4 Soit $0 \leq \varphi \in L^2(I, \mu)$ telle que $c(\varphi, \rho_0) < +\infty$. Alors pour tout $\nu = f\mu \in \mathcal{M}_1(I)$, on a :

$$(3.43) \quad \|\varphi(\nu - \mu)\|_{TV} \leq c(\varphi, \rho_0) \int_I \sqrt{\Gamma(f, f)} d\mu \leq 2c(\varphi, \rho_0) \sqrt{I(\nu|\mu)}.$$

D'une manière équivalente : pour toute fonction $g : I \rightarrow \mathbb{R}$ telle que $\|g\|_{Lip(\beta_\varphi)} \leq 1$ où $\beta_\varphi(x, y) := [\varphi(x) + \varphi(y)]1_{x \neq y}$, on a, pour toute mesure initiale $\nu \ll \mu$ et tous réels $t, r > 0$,

$$(3.44) \quad \mathbb{P}_\nu \left(\frac{1}{t} \int_0^t g(X_s) ds > \mu(g) + r \right) \leq \left\| \frac{d\nu}{d\mu} \right\|_{L^2(I, \mu)} \exp \left(- \frac{tr^2}{4c(\varphi, \rho_0)^2} \right).$$

L'inégalité de concentration Gaussienne (3.44) vient de (3.43) grâce à [95, Theorem 2.4] et à l'égalité $\|\varphi(\nu - \mu)\|_{TV} = \sup_{g: \|g\|_{Lip(\beta_\varphi)} \leq 1} \int_I g d(\nu - \mu)$ (cf. [91]).

On remarque que quand $\varphi = 1$ dans (3.43) alors $c(\varphi, \rho_0) = c_\delta$. Donc, dans ce cas particulier, l'inégalité (3.43) devient : pour toute mesure de probabilité μ de densité $f \in \mathcal{AC}(I)$,

$$(3.45) \quad \int_I |f - 1| d\mu \leq c_\delta \int_I \sqrt{\Gamma(f, f)} d\mu \leq 2c_\delta \sqrt{I(f\mu|\mu)}.$$

Guillin et al. [95, Theorem 3.1] ont prouvé que si l'inégalité de Poincaré a lieu, alors on a :

$$\int_I |f - 1| d\mu \leq \sqrt{2c_G I(\nu|\mu)}$$

avec une constante optimale $c_G \leq 2c_P$ (l'indice G indique l'inégalité équivalente pour la concentration Gaussienne). Inversement, si la dernière inégalité a lieu, alors on a $c_P \leq 2c_G$.

Les inégalités de concentration (3.42) et (3.44) ne contiennent pas la variance asymptotique de g définie par :

$$\sigma^2(g) := \lim_{t \rightarrow \infty} \frac{1}{t} \operatorname{Var}_{\mathbb{P}_\mu} \left(\int_0^t g(X_s) ds \right).$$

Cette variance joue un rôle fondamental dans le TCL. Ceci est traité dans l'inégalité de concentration de type Bernstein ci-après.

Corollaire 3.3.5 *Supposons que la constante c_δ dans (3.38) est finie.*

- (i) *Si la constante c_ρ dans (3.40) est finie, alors pour toute fonction g telle que $\|g\|_{Lip(\rho)} \leq 1$, on a, pour toute mesure initiale $\nu \ll \mu$ et pour tous réels $t, x > 0$,*

$$\mathbb{P}_\nu \left(\frac{1}{t} \int_0^t g(X_s) ds > \mu(g) + \sqrt{(2\sigma^2(g) + 4c_\rho^2 \min\{1, c_\delta \sqrt{x}\}) x} \right) \leq \left\| \frac{d\nu}{d\mu} \right\|_{L^2(I, \mu)} e^{-tx}.$$

- (ii) *Si la constante $c(\varphi, \rho_0)$ dans (3.31) est finie, alors pour toute fonction mesurable g telle que $|g(x) - g(y)| \leq \varphi(x) + \varphi(y)$, l'inégalité dans (i) est vraie en remplaçant c_ρ par $c(\varphi, \rho_0)$.*

Une description qualitative de la bornitude de l'opérateur de Poisson

Pour $g \in L_0^2(I, \mu)$, la solution G avec $\mu(G) = 0$ de l'équation de Poisson $-\mathcal{L}_2 G = g$, si elle existe, sera notée par $(-\mathcal{L})^{-1}g$.

On peut penser naturellement que lorsque φ est bornée mais tend vers 0 sur le bord ∂I , la norme Lipschitzienne $c(\varphi, \rho_0)$ peut être finie même si $c_\delta = +\infty$. À première vue, on pourrait penser faire de même pour $c_{Lip}(\rho, \rho_0)$ quand ρ' tend vers 0 sur le bord ∂I . Cependant ceci n'est pas possible.

Proposition 3.3.6 *Soient ρ, φ comme dans le Théorème 3.3.1, mais de plus bornées et soit $\varphi > 0$. Soit $\rho_0(x) = \int_c^x \frac{1}{\sqrt{a(y)}} dy$. Considérons les propriétés suivantes :*

- (i) $c_\rho = c_{Lip}(\rho, \rho_0) = \|(-\mathcal{L}_2)^{-1}\|_{C_{Lip(\rho),0} \rightarrow C_{Lip(\rho_0),0}} < +\infty$.
- (ii) $c(\varphi, \rho_0) = \sup_{g: |g| \leq \varphi} \|(-\mathcal{L}_2)^{-1}(g - \mu(g))\|_{Lip(\rho_0)} < +\infty$.
- (iii) $c_\delta = \sup_{|g| \leq 1} \|(-\mathcal{L}_2)^{-1}(g - \mu(g))\|_{Lip(\rho_0)} < +\infty$.
- (iv) *L'inégalité L^1 de Poincaré (3.34) a lieu, c.à.d., $c_{P,1} < +\infty$.*
- (v) *L'inégalité de transport-information suivante a lieu : il existe une certaine constante $c_G > 0$ telle que, pour $\nu = f\mu \in \mathcal{M}_1(I)$, on a $\int_I |f - 1| d\mu \leq \sqrt{2c_G I(\nu|\mu)}$.*
- (vi) *L'inégalité de Poincaré (3.27) a lieu, c.à.d., $c_P < +\infty$.*

Alors

1. *les propriétés (i), (ii), (iii) et (iv) sont équivalentes.*
2. *(iv) implique (v) et (v) et (vi) sont équivalentes.*
3. *Si $a(x) = 1$ et $b' \leq K$ (c.à.d. la courbure de Bakry-Emery est bornée inférieurement par $-K$), alors (vi) implique (iv) et alors (i), (ii), (iii), (iv), (v), et (vi) sont toutes équivalentes.*

3.3.3 Exemples

Mesure Gaussienne

Soit $I = \mathbb{R}$, $a(x) = 1$ et $b(x) = -x/\sigma^2$ où $\sigma > 0$. Alors $m'(x) = e^{-x^2/2\sigma^2}$ et $\mu = \mathcal{N}(0, \sigma^2)$ (la loi Gaussienne centrée de variance σ^2). Pour $\rho_0(x) = x$, on voit que :

$$c_{Lip}(\rho_0, \rho_0) = c_{\rho_0} = \sup_{x \in \mathbb{R}} e^{x^2/2\sigma^2} \int_x^\infty y e^{-y^2/2\sigma^2} dy = \sigma^2.$$

On a $c_P = c_{\rho_0} = \sigma^2$, qui est une égalité bien connue (voir [115]). L'inégalité de transport (3.41) devient une égalité pour $\nu = \mathcal{N}(m, \sigma^2)$.

On identifie la constante c_δ dans (3.38) par un calcul direct :

$$c_\delta = 2 \sup_{x \in \mathbb{R}} e^{x^2/2\sigma^2} \sqrt{2\pi} \sigma \mu([x, +\infty)) \mu((-\infty, x)) = \sqrt{\frac{\pi}{2}} \sigma.$$

D'un autre côté, on a $c_{cheeger} \geq \sqrt{\frac{\pi}{2}} \sigma$ comme on l'a déjà vu pour $A = \mathbb{R}^+$. Alors d'après les inégalités (3.36) et le Théorème 3.3.2, on obtient $c_{cheeger} = c_\delta = c_{P,1}$.

Mesure uniforme

Soit $I = [-D/2, D/2]$ où $D > 0$, $a(x) = 1$ et $b(x) = 0$. L'unique mesure invariante de probabilité μ est la mesure uniforme sur I . Comme $m'(x) = 1 = s'(x)$, on obtient

$$c_{\rho_0} = c_{\text{Lip}(\rho_0, \rho_0)} = \sup_{x \in [-D/2, D/2]} \int_x^{D/2} y dy = \frac{D^2}{8},$$

et la constante $c_\delta = c(\varphi, \rho_0)$ avec $\varphi = 1$ est donnée par

$$c_\delta = \sup_{x \in [-D/2, D/2]} 2D\mu([-D/2, x])\mu([x, D/2]) = \frac{D}{2}.$$

Comme $c_{\text{cheeger}} \geq D/2$ (ainsi que nous l'avons déjà vu pour $A = [0, D/2]$), on obtient $c_{\text{cheeger}} = D/2 = c_\delta = c_{P,1}$ par l'inégalité (3.36) et le théorème 3.3.2.

Mesure exponentielle sur \mathbb{R}^+

Soit $I = \mathbb{R}^+ = [0, +\infty)$, $a(x) = 1$ et $b(x) = -\lambda$ où $\lambda > 0$. Alors $m'(x) = e^{-\lambda x} = 1/s'(x)$, $\rho_0(x) = x$ et μ est la distribution exponentielle de paramètre λ . Il est facile de voir que $c_{\rho_0} = c_{\text{Lip}(\rho_0, \rho_0)} = +\infty$: pas de trou spectral pour la norme ρ_0 -Lipschitzienne. En effet, l'inégalité de transport-information (3.41) est fausse pour $\rho = \rho_0$. Par le Théorème 3.3.2, on trouve

$$c_{P,1} = c_\delta = 2 \sup_{x \geq 0} \frac{1}{\lambda} e^{\lambda x} \mu(0, x) \mu(x, +\infty) = 2 \sup_{x \geq 0} \frac{1}{\lambda} \mu(0, x) = \frac{2}{\lambda}.$$

Cependant, on a $c_{\text{cheeger}} = \frac{1}{\lambda}$ par Bobkov-Houdré [30]. Ceci, ainsi que l'étude de la mesure Gaussienne précédente, montre que les deux inégalités dans (3.36) sont optimales. On a aussi l'inégalité de transport-information (3.45), qui peut-être écrite selon :

$$\|\nu - \mu\|_{TV} \leq \frac{4}{\lambda} \sqrt{I(\nu|\mu)}, \quad \forall \nu.$$

Elle est optimale. Pour ce modèle il est bien connu que $c_P = 4/\lambda^2$ ([115]). L'inégalité précédente est la même que celle prouvée dans [95, Theorem 3.1] (à partir de l'inégalité de Poincaré).

Mesure Log-concave sur \mathbb{R}

Soit $I = \mathbb{R}$, $a(x) = 1$ et $b(x) = -V'(x)$ où V est C^2 , strictement convexe sur \mathbb{R} telle que $V(0) = 0$ et $\int_{\mathbb{R}} e^{-V} dx < +\infty$. Alors $m'(x) = e^{-V(x)}$, $s'(x) = e^{V(x)}$ et $\rho_0(x) = x$.

Soit $\rho(x) = V'(x)$, qui est μ -intégrable et $\mu(\rho) = 0$. On a

$$c_\rho = c_{\text{Lip}(\rho, \rho_0)} = \sup_{x \in \mathbb{R}} e^{V(x)} \int_x^{+\infty} V'(y) e^{-V(y)} dy = \sup_{x \geq 0} e^{V(x)} e^{-V(x)} = 1.$$

Ainsi en supposant que $\int V'^2 e^{-V} dx < +\infty$, on obtient l'inégalité de transport-information (3.41) et l'inégalité de concentration Gaussienne (3.42). Par exemple, pour toute fonction $g \in C^1(\mathbb{R})$ telle que, $|g'| \leq V''$ on a, pour toute mesure initiale $\nu \ll \mu$ et tous réels $t, r > 0$,

$$(3.46) \quad \mathbb{P}_\nu \left(\frac{1}{t} \int_0^t g(X_s) ds > \mu(g) + r \right) \leq \left\| \frac{d\nu}{d\mu} \right\|_{L^2(I, \mu)} \exp \left(-\frac{tr^2}{4} \right).$$

En outre, pour toute fonction positive $\varphi \leq M(1 + |V'|)$, il est facile de voir que $c(\varphi, \rho_0) < +\infty$. Dès lors, l'inégalité de transport-information (3.43) est satisfaite.

Il est très intéressant de généraliser ce résultat aux mesures log-concave sur \mathbb{R}^d . Voir Bobkov-Ledoux [31] pour certains résultats dans cette direction.

Diffusion de Jacobi

Soit $I =]0, 1[$, $a(x) = x(1-x)$ et $b(x) = -x + 1/2$, alors $\mu(x) = 1/(\pi \sqrt{x(1-x)})$, voir [65]. Pour $\rho_0(x) = \frac{\pi}{2} + \text{Arcsin}(2x-1)$, on voit que

$$c_{\text{Lip}(\rho_0, \rho_0)} = c_{\rho_0} = \sup_{x \in]0, 1[} \left(\frac{\pi^2}{8} - \frac{1}{2} \text{Arcsin}^2(2x-1) \right) = \frac{\pi^2}{8}.$$

Par le calcul on identifie la constante c_δ dans (3.38) comme

$$c_\delta = \frac{2}{\pi} \sup_{x \in]0,1[} \left(\frac{\pi^2}{4} - \text{Arcsin}^2(2x-1) \right) = \frac{\pi}{2}.$$

En utilisant (3.37), voir Bobkov-Houdré [30], on obtient $c_{cheeger} = \frac{\pi}{2}$, ainsi on trouve $c_{cheeger} = c_{P,1} = c_\delta = \frac{\pi}{2}$ d'après le Théorème 3.3.2.

Processus de branchement continu

Soit $I =]0, +\infty[$, $a(x) = 2x$ et $b(x) = -2x + 1$, alors $\mu(x) = \frac{1}{\sqrt{\pi}} \frac{e^{-x}}{\sqrt{x}}$. Ce processus apparaît comme une diffusion limite de certains processus de branchement à temps discret, voir [118]. Pour $\rho_0(x) = \sqrt{2x}$, on voit que

$$c_{Lip}(\rho_0, \rho_0) = c_{\rho_0} = \sup_{x \in \mathbb{R}^+} \left(1 - \frac{e^x}{\sqrt{\pi}} \int_x^\infty \frac{e^{-y}}{\sqrt{y}} dy \right) = 1.$$

Exemple 3.3.7 (voir aussi l'exemple 1.4.2 dans [168])

Soit $I = \mathbb{R}^+$, $a(x) = (1+x)^\alpha$ avec $\alpha > 1$, et $b(x) = 0$ alors $\mu(x) = \frac{\alpha-1}{(1+x)^\alpha}$. Pour $\alpha > 2$ et $\rho_0(x) = \frac{2}{\alpha-2}(1 - (1+x)^{-\frac{\alpha-2}{2}})$, on voit que :

$$c_{Lip}(\rho_0, \rho_0) = c_{\rho_0} = \frac{4}{(\alpha-2)(3\alpha-4)} \sup_{x \in \mathbb{R}^+} (1+x)^{-\frac{\alpha-2}{2}} \left(1 - (1+x)^{-\frac{\alpha-2}{2}} \right) = \frac{1}{(\alpha-2)(3\alpha-4)}.$$

Par le calcul, on identifie la constante c_δ dans (3.38) :

$$c_\delta = \frac{2}{\alpha-1} \sup_{x \in \mathbb{R}^+} (1+x)^{-\frac{\alpha-2}{2}} \left(1 - (1+x)^{-(\alpha-1)} \right) = \frac{4}{3\alpha-4} \left(\frac{\alpha-2}{3\alpha-4} \right)^{\frac{\alpha-2}{2(\alpha-1)}}.$$

D'après le Théorème 3.3.2, on a $c_{P,1} = c_\delta$. Cependant, en utilisant (3.37), on obtient $c_{cheeger} = \frac{1}{\alpha-1} \left(\frac{1}{2} \right)^{\frac{\alpha-2}{2(\alpha-1)}}$.

Problèmes et perspectives

Nous continuons à nous intéresser aux inégalités de transport car un de leurs intérêts principaux est leur lien avec la concentration de la mesure.

1. Notre objectif est d'étendre les inégalités de transport que nous avons prouvées dans le cas de diffusion, à des équations différentielles stochastiques rétrogrades. Une généralisation a déjà été faite par Saussereau [155] dans le cas des équations différentielles stochastiques conduite par un mouvement brownien fractionnaire.
2. Nous souhaitons étudier quelques inégalités fonctionnelles, comme les inégalités de Poincaré, les inégalités de log-Sobolev et les inégalités de transport pour les modèles de chaînes de Markov bifurcantes, sur un arbre binaire. Le fait de ne plus considérer un arbre binaire pour la population de cellules va modifier qualitativement les résultats, dès que les transitions vont dépendre du nombre de cellules filles. Delmas et Marsalle [60] ont considéré le cas où une cellule peut avoir 0,1 ou 2 enfants en forçant la généalogie sous-jacente à la population de cellules à être un processus de Galton-Watson. Nous envisageons de traiter ce cas aussi.
3. L'extension de l'étude de l'évolution d'un système de particules dont la généalogie est donnée par un arbre de Galton-Watson continu est proposée dans [13]. Les particules se déplacent de manière indépendante en fonction d'un processus de Markov et lorsque se produit un événement de branchement, les positions des descendants dépendent de la position de la mère et du nombre de descendants. Une loi des grands nombres pour la mesure empirique des individus vivant au temps t est obtenue. Nous souhaitons obtenir des inégalités de déviations dans ce cas.
4. Nous envisageons de regarder les inégalités de transport sur les graphes.

Chapitre 4

Déviations modérées pour des variables aléatoires dépendantes

Dans ce chapitre, nous présentons succinctement des résultats de déviations modérées pour certaines variables aléatoires dépendantes pour les chaînes de Markov, pour des différences de martingales, pour des processus autorégressifs et pour la densité spectrale d'un processus de moyennes mobiles. Ces résultats sont ceux des articles [T-D2], [T-D3], [H-D6], [H-D9], [T-D17] et [T-D19].

4.1 Introduction

Le but est ici d'étendre à des situations de dépendances, les conditions nécessaires et suffisantes obtenues par Ledoux-Arcones [113, 8] dans le cas de sommes de variables aléatoires i.i.d. basées sur le comportement de la queue de distribution de ces variables par rapport à la vitesse des déviations modérées.

Rappelons que Borovkov et Mogulskii [38] ont considéré le PDM pour une suite de v.a.i.i.d. (X_i) à valeurs dans un espace de Banach, sous la condition que $\mathbb{E}e^{\delta|X_1|} < +\infty$, pour un certain $\delta > 0$. Baldi [12] a obtenu le PDM trajectorien sous la même condition.

Dans le cas où la vitesse b_n du PDM est égale à n^α avec $\frac{1}{2} < \alpha < 1$, Chen [46] a trouvé une condition nécessaire et suffisante pour le PDM dans un espace Banach et a obtenu la borne inférieure pour une suite b_n générale sous des conditions très faibles.

En utilisant les techniques d'isopérimétrie, Ledoux [113] a obtenu une condition nécessaire et suffisante dans le cas général [c'est-à-dire pour n'importe quelle vitesse b_n du PDM], prolongeant ainsi les travaux de Chen [46].

Plus précisément, il a prouvé l'équivalence entre le PDM de vitesse b_n et la condition de queue exponentielle suivante :

$$(4.1) \quad \exists R > 0, \text{ tel que } \forall u > 0, \limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \left(n\mathbb{P}(|X_1| > ub_n) \right) \leq -\frac{u^2}{R},$$

sous réserve que b_n satisfasse une condition technique : il existe $A > 1$, $0 < \delta < 1$ tels que $\forall n, k \geq 1$, $b_{nk} \leq Ak^{1-\delta}b_n$, ce qui signifie que b_n ne peut être trop près de n , l'échelle des grandes déviations.

Pour des déviations modérées indépendantes de la vitesse b_n considérée, Jiang a montré qu'un moment exponentiel est nécessaire.

Les résultats de Ledoux [113] ont été étendus aux processus empiriques fonctionnels (dans le cadre de la statistique non paramétrique) par Wu [171]. D'autres développements ont été obtenues par Dembo et Zajic [65]. Arcones [8] a obtenu le PDM de type fonctionnel, sans l'hypothèse technique sur b_n de Ledoux [113], mais avec (4.1) remplacée par

$$(4.2) \quad \limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \left(n\mathbb{P}(|X_1| > b_n) \right) = -\infty.$$

Sous cette condition et dans le cas i.i.d. voir aussi les travaux de Eichelsbacher et Löwe [75]. La question de savoir comment étendre le principe de déviations modérées à des situations de dépendance a récemment fait l'objet de beaucoup d'attention et de travaux remarquables.

4.2 Chaînes de Markov

Soit $(X_n)_{n \geq 0}$ une chaîne de Markov, à valeurs dans un espace d'état mesurable (E, \mathcal{E}) , ergodique, de probabilité de transition P et de mesure invariante π . Soit $\mathcal{M}(E)$ l'espace des mesures signées à variations finies sur (E, \mathcal{E}) muni de la norme de la variation totale $\|\cdot\|_{var}$. Pour une mesure de probabilité μ sur (E, \mathcal{E}) , soit \mathbb{P}_μ la mesure de probabilité Markovienne sur $(E^{\mathbb{N}}, \mathcal{E}^{\otimes \mathbb{N}})$ déterminée par la probabilité de transition P et la mesure initiale μ . $\{X_j\}$ représente alors la suite des coordonnées sur $E^{\mathbb{N}}$. On définit les mesures empiriques M_n , $n \geq 1$

$$M_n := \frac{1}{b_n} \sum_{j=0}^{n-1} (\delta_{X_j} - \pi)$$

qui sont des éléments aléatoires de $\mathcal{M}(E)$. En statistique non paramétrique, on a besoin d'estimations uniformes de $M_n(f)$ sur une classe de fonctions \mathcal{F} .

Si $b_n = \sqrt{n}$, il s'agit de l'étude du TCL pour lequel nous renvoyons à Meyn-Tweedie [129]. Lorsque $b_n = n$, le problème devient celui des grandes déviations, amplement étudié dans la littérature : voir Donsker-Varadhan [68] pour les premiers travaux, jusqu'à Wu [178] pour les derniers développements.

Notre cadre d'étude sera celui où la vitesse b_n (supposée strictement croissante) vérifie

$$(4.3) \quad \sqrt{n} \ll b_n \ll n.$$

Notre but est ici d'introduire l'analogue de conditions du cas i.i.d. en termes de jeu entre la vitesse b_n , l'ergodicité de la chaîne de Markov, i.e. formellement une suite r_n tendant vers 0 et une fonction V telles que pour tout x

$$(4.4) \quad \|P^n(x, \cdot) - \pi\|_{var} \leq V(x)r_n,$$

et le caractère non borné de f . Nous étendons aussi lorsque cela est possible, en vue d'applications aux processus à temps continu, les déviations modérées au cas trajectorien et uniformément sur des classes de fonctions. Des applications aux lois du logarithme itéré ainsi que des extensions du TCL sont également obtenues.

Avant d'énoncer les principaux résultats obtenus dans cette partie, il est nécessaire d'introduire, au moins formellement, la méthode dite de *régénération de la chaîne éclatée* qui constitue la base de notre approche. Pour une approche complète de cette méthode, on pourra consulter Nummelin [136, 137], Athreya-Ney [9], Meyn-Tweedie [129] ou Chen [46]. Cette méthode est basée sur le fait que si nous considérons la chaîne de Markov (X_n) π -irréductible, alors il y a existence d'*ensembles petits* : $\exists m \in \mathbb{N}^*$, $\exists C \subset E$, $\nu \in \mathcal{P}(E)$, $\nu(C) = 1$, tels que $\forall x \in C$

$$(4.5) \quad P^m(x, \cdot) \geq \varepsilon \nu(\cdot).$$

Lorsqu'il existe C tel que $\forall x, y \in C$, $P(x, \cdot) = P(y, \cdot)$, C est nommé *atome*.

La méthode de *régénération de la chaîne éclatée* se base sur l'existence d'un ensemble petit, disons pour $m = 1$ pour simplifier, pour construire une chaîne de Markov sur $E \times \{0, 1\}$, soit (X_n, Y_n) dont la première marginale a la même loi que (X_n) mais qui a l'avantage d'avoir $C \times \{1\}$ comme atome : lorsque la chaîne n'est pas dans C elle évolue normalement avec la transition P , et $Y_n = 0$, mais lorsque la chaîne rentre dans C , on tire une Bernoulli indépendante de paramètre ε pour la deuxième composante et si $Y_n = 1$, X_{n+1} est tiré avec la probabilité ν sinon avec le noyau résiduel. Ainsi à chaque fois que $Y_n = 1$, la chaîne évolue selon ν et donc "oublie" le point précédent. La construction pour $m > 1$ est "identique" en raisonnant sur P^m .

De par la définition d'un atome, on voit que chaque passage dans un atome "régénère" la chaîne permettant ainsi de créer des blocs entre ces retours à l'atome qui sont indépendants si $m = 1$ et 1-dépendants sinon. Ainsi, en considérant pour simplifier $m = 1$, en notant (σ_k) les temps de passages successifs dans l'atome, $i(n)$ le nombre de passages dans l'atome avant n et $l(n)$ le dernier temps de passage dans l'atome avant n , on a (en supposant $\pi(f) = 0$)

$$(4.6) \quad M_n(f) = \frac{1}{b_n} \sum_{k=0}^{\sigma_1} f(X_k) + \frac{1}{b_n} \sum_{j=1}^{i(n)-1} \sum_{k=\sigma_j}^{\sigma_{j+1}} f(X_k) + \frac{1}{b_n} \sum_{k=l(n)}^{n-1} f(X_k).$$

Les premier et dernier termes s'avèreront négligeables à l'échelle des déviations modérées, ce qui fait que nous revenons à l'étude des déviations modérées d'une somme aléatoire de v.a. indépendantes (ou 1-dépendantes). L'ergodicité que nous imposons permettra de transformer cette somme aléatoire en une somme déterministe.

Nous supposons pour faire simple que cette chaîne possède un *atome*, i.e. $\exists \alpha \in \mathcal{E}$, une mesure de probabilité ν sur E tels que

$$\forall x \in \alpha, \quad P(x, \cdot) = \nu(\cdot)$$

et α est ainsi appelé *atome*. On remarquera que, dans le cas où E est discret, chaque état est un *atome*. Le cas markovien a été extensivement étudié sous des hypothèses de plus en plus générales et successivement moins restrictives (voir Wu [171, 172], Puhalskii [147], Chen [47] pour des références pertinentes), et récemment par de Acosta [53] et Chen [48] pour la borne inférieure sous la seule condition de l'ergodicité de degré 2, et de Acosta-Chen [54] et Chen [48] pour la borne supérieure sous l'ergodicité géométrique et une condition de régularité de la mesure initiale, dont l'objectif principal est le principe de déviations modérées pour les fonctionnelles bornées à valeurs dans un espace de Banach séparable de M_n et pour la mesure empirique M_n . Guillin [93] a obtenu le principe de déviations modérées uniforme en temps pour les processus empiriques fonctionnels. Ces travaux ont fondamentalement motivé notre travail. Nous améliorons et étendons ces résultats.

Précisons quelques notations. Etant donnée une classe de fonctions réelles et mesurables \mathcal{F} telle que $\forall f \in \mathcal{F}$ on a $\pi(f) = 0$, $f \in L_2(\pi)$, soit $l_\infty(\mathcal{F})$ l'espace de toutes les fonctions réelles bornées sur \mathcal{F} , muni de la norme $\|F\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} |F(f)|$, qui est un espace de Banach non séparable quand \mathcal{F} est infini. A toute mesure $\nu \in \mathcal{M}(E)$ correspond un élément $\nu^{\mathcal{F}}$ dans $l_\infty(\mathcal{F})$ donné par $\nu^{\mathcal{F}} = \int f d\nu$, $\forall f \in \mathcal{F}$. Afin d'éviter les problèmes de mesurabilité, on suppose que \mathcal{F} est dénombrable, ou que le processus $\{M_n(f), f \in \mathcal{F}\}$ est séparable au sens de Doob pour tout n . Nous nous intéressons ici au comportement asymptotique de $\mathbb{P}_\mu((M_n)^{\mathcal{F}} \in \cdot)$.

Théorème 4.2.1 *Supposons que (\mathcal{F}, d_2) est totalement bornée et que $(M_n)^{\mathcal{F}} \rightarrow 0$ en probabilité dans $l_\infty(\mathcal{F})$. Supposons que*

$$(H1) \quad \limsup_{n \rightarrow +\infty} \frac{n}{b_n^2} \log(n \mathbb{P}_\nu(\tau \geq b_n)) = -\infty;$$

$$(H2) \quad \limsup_{n \rightarrow +\infty} \frac{n}{b_n^2} \log \left(n \mathbb{P}_\nu \left(\sum_{k=0}^{\tau} \|f(X_k)\|_{\mathcal{F}} \geq b_n \right) \right) = -\infty.$$

Alors, pour toute mesure de probabilité μ sur (E, \mathcal{E}) vérifiant

$$(4.7) \quad \limsup_{n \rightarrow +\infty} \frac{n}{b_n^2} \log \mathbb{P}_\mu \left(\sum_{k=0}^{\tau} \|f(X_k)\|_{\mathcal{F}} \geq b_n \right) = -\infty.$$

$\mathbb{P}_\mu((M_n)^{\mathcal{F}} \in \cdot)$ satisfait le PDM sur $l_\infty(\mathcal{F})$ de vitesse b_n^2/n et de bonne fonction de taux $J_{\mathcal{F}}$ donnée par

$$J_{\mathcal{F}}(F) = \sup \left\{ J_{(f_1, \dots, f_m)}(F(f_1), \dots, F(f_m)); f_1, \dots, f_m \in \mathcal{F}, m \geq 1 \right\},$$

où $J_{(f_1, \dots, f_m)}$ est donnée par $J_f(x) = \sup_{\xi \in \mathbb{R}^m} \left[\langle x, \xi \rangle - \frac{1}{2} \sigma^2(\langle f, \xi \rangle) \right]$.

Supposons par ailleurs que $\sum_{k=1}^{\infty} \langle \xi, f \rangle P^k \langle f, \xi \rangle \in L_1(\pi)$ pour tout $\xi \in \mathbb{R}^m$, alors

$$\sigma^2(\langle f, \xi \rangle) = \int \langle f, \xi \rangle^2 d\pi + 2 \int \sum_{k=1}^{\infty} \langle f, \xi \rangle P^k \langle f, \xi \rangle d\pi.$$

Avant de décrire brièvement la preuve, faisons quelques commentaires sur nos hypothèses.

Notons qu'en présence d'atome, l'ergodicité géométrique est équivalente à $\exists \delta > 0$ tel que $\mathbb{E}_\nu(e^{\delta \tau}) < \infty$. La condition **(H1)** est ainsi strictement plus faible que l'ergodicité géométrique imposée dans les travaux de de Acosta-Chen [54]. De plus **(H1)** peut-être explicitement donnée. Par exemple, dans le cas particulier $b_n = n^{1/p}$ avec $1 < p < 2$, alors **(H1)** est vraie, si

$$(4.8) \quad \exists \delta > 0 \text{ tel que } \mathbb{E}_\nu(e^{\delta \tau^{2-p}}) < \infty.$$

Remarquons également que nous considérons ici le cas non séparable du processus empirique fonctionnel et des fonctions non bornées. Ces cas ne sont pas étudiés par de Acosta-Chen. À leur crédit, notons toutefois qu'ils ne supposent pas l'existence d'un atome et leur seule hypothèse est l'ergodicité géométrique.

Toujours dans le contexte $b_n = n^{\frac{1}{p}}$ avec $1 < p < 2$, en suivant Nummelin-Tweedie [139] et Nummelin-Tuominen [138] (ou Meyn-Tweedie [129] pour un examen complet) on peut voir que la condition (4.8) est équivalente à l'ergodicité sous-géométrique suivante : $\exists r > 1$ tel que pour π -presque sûrement x

$$\sum_{n=1}^{\infty} r^{n^{2-p}} \|P^n(x, \cdot) - \pi\|_{var} < \infty,$$

condition qui est plus forte que l'ergodicité de degré 2 (voir Chen [47]) mais plus faible que l'ergodicité géométrique. Une telle affirmation implique en particulier que (4.8) est valable indépendamment du choix de l'atome et donc **(H1)** aussi dans ce contexte.

Nous ne sommes pas en mesure d'obtenir l'indépendance de la condition de récurrence **(H1)** de l'atome, ni sa caractérisation au moyen d'un certain type d'ergodicité pour un b_n général, mais heureusement, nos résultats sont prouvés si **(H1)** et **(H2)** sont satisfaites pour un certain atome et puis pour tout atome.

Sous **(H2)**, la condition (4.7) est vérifiée, par exemple, par la mesure invariante π de la chaîne de Markov et par la mesure de Dirac δ_x pour π -presque tout $x \in E$.

La preuve se décompose en deux étapes : utilisation de la formule (4.6) pour établir des déviations modérées avec une condition de type Ledoux-Arcones sur les excursions de la fonctionnelle entre 2 temps de retour à l'atome, puis utilisation de la condition de dérive pour contrôler ces excursions par des quantités reposant sur des caractéristiques de la chaîne initiale : l'introduction d'une chaîne intermédiaire, ne régénérant pas, est cruciale pour le contrôle des moments des excursions, les conditions de dérive permettant, par le biais de la formule de Dynkin, de majorer les moments des excursions de cette chaîne.

Notons pour conclure que dans [T-D2], publié dans *Stochastic Processes and Their Applications* avec A. GUILLIN, nous donnons des exemples, dans le cas de chaîne de Markov à valeurs dans un espace d'état dénombrable, de classes de fonctions vérifiant les conditions de ces théorèmes. Pour l'extension au cadre général d'une chaîne de Markov sans atome voir [T-D17] publié aux *CRAS* avec A. GUILLIN.

4.3 Différences de martingales et suites ϕ -mélangeantes

Nous présentons dans cette partie notre travail [T-D4], publié dans *Stochastic and Stochastic Reports*.

Soit $(m_k)_{k \geq 1}$ une suite de différences de martingales à valeurs dans \mathbb{R}^d définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, par rapport à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$.

Soit $M_0 = 0$, $M_n = \sum_{k=1}^n m_k$, $\forall n \in \mathbb{N}^*$ et on note par $\langle M \rangle_n$ le processus de la variation quadratique de la martingale (M_n) donné par $\langle M \rangle_n = \sum_{k=1}^n \mathbb{E}_{k-1}(m_k m_k^*)$.

Ici, \mathbb{E}_{k-1} et \mathbb{P}_{k-1} désignent respectivement l'espérance conditionnelle et la probabilité conditionnelle sachant \mathcal{F}_{k-1} .

Pour de nombreuses applications dans les statistiques, on cherche le comportement limite de $\mathbb{P}\left(\frac{1}{b_n}M_n \in \cdot\right)$. Des bornes uniformes et non uniformes sur la convergence dans le TCL de M_n/\sqrt{n} ont été obtenues par plusieurs auteurs, voir Račkauskas [148], [149], Grama [83], Grama et Haeusler [92]. Pour la preuve du TCL et pour des informations générales sur le TCL de martingales, on se réfère à Hall et Heyde [98], Liptser et Shiryaev [119] ou Jacod et Shiryaev [106].

Notre objectif dans ce travail est de donner le comportement asymptotique des fonctionnelles associées au principe d'invariance de Donsker $Z_n(t) = M_{[nt]}/b_n$, $t \in [0, 1]$, dans $(\mathbb{D}[0, 1], \mathbb{R}^d)$, l'espace des fonctions càdlàg définies sur $[0, 1]$ à valeurs dans \mathbb{R}^d muni de la topologie de Skorohod et de la σ -algèbre \mathcal{B} .

L'étude du principe de déviations modérées des martingales est récent : voir Puhalskii [145], Račkauskas [149], Dembo [61], Gao [84] et Worms [170]. Račkauskas [149] a obtenu la borne supérieure des déviations modérées pour une suite réelle de différences de martingales bornées $Z_n(1)$ sous des conditions très restrictives. Sa preuve est basée sur des estimations précises du taux de convergence de M_n/\sqrt{n} vers la loi normale.

Dans le cas de différences de martingales bornées à valeurs dans \mathbb{R}^d , Dembo [61], en utilisant la méthode des cumulants de Puhalskii [145], a donné le PDM pour $Z_n(\cdot)$ dans $\mathbb{D}[0, \infty)$, muni de la topologie localement uniforme, sous la condition suivante : pour une certaine matrice symétrique semi-définie positive Q , on a

$$\forall \delta > 0 \quad \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P} \left(\left\| \frac{\langle M \rangle_t}{t} - Q \right\| > \delta \right) < 0.$$

Puhalskii [145] a établi des conditions pour obtenir le PGD sous la topologie de Skorohod en termes de convergence de leurs caractéristiques prévisibles. Gao [84] a discuté le PDM pour une suite des différences de martingales sous une condition d'intégrabilité conditionnelle de la martingale, avec quelques applications aux suites mélangeantes.

Worms [170, 169] donne des critères pour le PDM d'une martingale vectorielle (et d'une certaine classe de séries de régression) avec une normalisation déterministe ou une autonormalisation. Certains de ses résultats sont basés sur la méthode des cumulants développée par Puhalskii [145].

Notre objectif est de prouver un théorème de type Chen-Ledoux pour le PDM d'une suite de différences de martingales. Notre principal résultat est le suivant

Théorème 4.3.1 *Soit $(b_n = b(n))$ une suite de nombre réels satisfaisant (4.3), telle que $c(n) := n/b_n$ est croissante. Nous définissons la fonction réciproque $c^{-1}(t)$ par $c^{-1}(t) := \inf \{n \in \mathbb{N} : c(n) \geq t\}$. Sous les hypothèses suivantes :*

(H1) *il existe une matrice Q symétrique semi-définie positive, telle que $\forall \delta > 0$,*

$$\limsup_{n \rightarrow +\infty} \frac{n}{b_n^2} \log \mathbb{P} \left(\left\| \frac{\langle M \rangle_n}{n} - Q \right\| > \delta \right) = -\infty ;$$

(H2) $\limsup_{n \rightarrow +\infty} \frac{n}{b_n^2} \log \left(n \operatorname{ess\,sup}_{1 \leq k \leq c^{-1}(b(n+1))} \mathbb{P}_{k-1} \left(|m_k| > b_n \right) \right) = -\infty ;$

(H3) $\forall a > 0$, and $\forall \delta > 0$

$$\limsup_{n \rightarrow +\infty} \frac{n}{b_n^2} \log \mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n \mathbb{E}_{k-1} \left(|m_k|^2 \mathbf{1}_{\{|m_k| \geq a \frac{n}{b_n}\}} \right) \geq \delta \right) = -\infty ;$$

la suite $Z_n(\cdot)$ satisfait le PDM sur $\mathbb{D}([0, 1], \mathbb{R}^d)$ de vitesse b_n^2/n et de bonne fonction de taux donnée par

$$(4.9) \quad I(\phi) = \begin{cases} \int_0^1 \Lambda^*(\phi'(t)) dt & \text{si } \phi \in \mathcal{AC}_0([0, 1]) \\ +\infty & \text{sinon,} \end{cases}$$

où Λ^* est donnée par $\Lambda^*(v) = \sup_{\lambda \in \mathbb{R}^d} (\lambda^* v - \frac{1}{2} \lambda^* Q \lambda)$ ($= \frac{1}{2} v^* Q^{-1} v$, si Q est inversible), et

$$\mathcal{AC}_0([0, 1]) = \left\{ \phi : [0, 1] \rightarrow \mathbb{R}^d, \text{ absolument continue, avec } \phi(0) = 0 \right\}.$$

Notre méthode de la preuve est la suivante :

- (1) pour les petit sauts de $Z_n(\cdot)$, on applique les résultats généraux de Puhalskii [145] ;
- (2) En suivant la méthode de Ledoux [113] et Arcone [8], on montre que la partie des grands sauts de $Z_n(\cdot)$ est négligeable au sens des déviations modérées.

Plusieurs difficultés techniques apparaissent dans la seconde étape, quand on passe du cas i.i.d. de Ledoux-Arcones au cas d'une martingale générale ici.

Notons que même pour des martingales (M_n) à sauts stationnaires ergodiques bornés avec $\mathbb{E}(m_1)^2 = 1$, la condition **(H1)** est indispensable pour le PDM. Un contre exemple est donné dans [61].

En contraste avec le PDM, notons que le PGD de vitesse n peut ne pas avoir lieu pour M_n/n même quand (M_n) est une martingale discrète à valeurs réelle avec des sauts bornés et indépendants telles que $\mathbb{E}(m_1^2) = 1$ voir [T-D4].

Si (M_t) est continue, alors **(H1)** (sans **(H2)**, **(H3)**) seule est suffisante pour la preuve du PDM de $Z_n(\cdot)$ sur l'espace des fonctions continues, avec une vitesse b_n^2/n et une bonne fonction de taux donnée dans (4.9). C'est une conséquence du lemme d'approximation [64, Theorem 4.2.13], de la représentation de Skorohod des martingales [150, Theorem 1.6, Chapter V] et du théorème de Schilder [64, Theorem 5.2.3].

Nous présentons ensuite un exemple de notre résultat appliqué à une certaine classe de suite ϕ -mélangeantes. Pour d'autres applications aux martingales stationnaires ou aux processus de Markov satisfaisant la condition de récurrence de Doeblin, voir [T-D4].

Soit $(X_n, n \in \mathbb{Z})$ une suite de variables aléatoires à valeurs dans un espace polonais (E, \mathcal{E}) définies sur un certain espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. On définit les σ -algèbres $\mathcal{F}_n^m = \sigma(X_k, m \leq k \leq n)$, $\mathcal{F}_n = \sigma(X_k, k \leq n)$ et $\mathcal{F}^m = \sigma(X_k, k \geq m)$. Posons

$$\phi(n) = \sup \left(|\mathbb{P}(B/A) - \mathbb{P}(B)| ; A \in \mathcal{F}_k \text{ avec } \mathbb{P}(A) > 0, B \in \mathcal{F}^{k+n}, k \in \mathbb{Z} \right).$$

La suite $(X_n, n \in \mathbb{Z})$ est dite ϕ -mélangeante si $\phi(n) \rightarrow 0$ quand $n \rightarrow +\infty$.

Supposons que le processus $(X_n, n \in \mathbb{Z})$ est stationnaire et ϕ -mélangeant, avec ϕ telle que :

$$(4.10) \quad \sum_{n=1}^{+\infty} \sqrt{\phi(n)} < +\infty, \quad \text{et} \quad \exists C \text{ une constante positive telle que } \mathbb{P}_0(X_1 \in \cdot) \leq C \mu(\cdot),$$

Corollaire 4.3.2 Soit f une fonction réelle mesurable dans $L^4(\mu)$ telle que $\mu(f) = 0$. Si

$$(4.11) \quad \limsup_{n \rightarrow +\infty} \frac{n}{b_n^2} \log(n \mathbb{P}(|f(X_0)| > ab_n)) = -\infty, \quad \forall a > 0,$$

alors $(\sum_{k=0}^{[nt]} f(X_k)/b_n)_{t \in [0, 1]}$ satisfait le PDM sur $\mathbb{D}[0, 1]$, de vitesse b_n^2/n et de bonne fonction de taux

$$J^f(\gamma) = \begin{cases} \frac{1}{2\sigma^2(f)} \int_0^1 \gamma'(t)^2 dt & \text{si } \gamma \in \mathcal{AC}_0([0, 1]) \text{ et } \sigma^2(f) > 0 \\ +\infty & \text{sinon,} \end{cases}$$

$\sigma^2(f)$ étant donnée par $\sigma^2(f) = \mathbb{E}(f(X_0)^2) + 2 \sum_{k=1}^{+\infty} \mathbb{E}(f(X_0)f(X_k))$.

L'idée de base pour ce résultat est la représentation des incréments stationnaires du processus en termes d'incrémentes de différences de martingales, plus d'autres termes qui seront négligeables au sens des grandes déviations. Cette idée est due à Gordin [89], qui l'a utilisée pour prouver le TCL pour des suites stationnaires ergodiques. Elle s'est largement développée ensuite (voir [98]).

Comparée à la convergence hyper-exponentielle de $\phi(n)$ vers 0 (c.à.d. $\phi(n)e^{\lambda n} \rightarrow 0, \forall \lambda > 0$) requise pour les grandes déviations dans le travail de Bryc [26], la condition (4.10) pour le principe de déviations modérées est beaucoup plus faible, c'est une condition très ancienne suffisante pour le TCL, (voir Corollary 5.5 dans [98]).

La condition (4.11), étant la meilleure possible, comme on l'a déjà vu dans le cas de v.a.i.i.d., est plus faible que la bornitude de f , imposée dans le travail de Gao [84], mais notre condition (4.10) est plus forte que celle imposée dans son travail.

4.4 Estimation de la densité spectrale d'un processus à moyenne mobile

Cette section présente les résultats obtenus en collaboration avec A. GUILLIN dans [T-D3] et A. GUILLIN et L. WU dans [H-D6], et nous nous référons également à [T-D19] pour une approche plus générale mais sous des hypothèses abstraites.

Le problème est ici d'étudier le comportement asymptotique de différents estimateurs associés au processus à moyenne mobile (X_n) , à valeurs réelles pour simplifier, défini par

$$X_n = \sum_{j=-\infty}^{+\infty} a_{j-n} \xi_j = \sum_{j=-\infty}^{+\infty} a_j \xi_{j+n}, \quad \forall n \in \mathbb{Z}$$

où $(\xi_i)_{i \in \mathbb{Z}}$ est une suite de variables réelles i.i.d. centrées de carré intégrable, de loi commune $\mathcal{L}(\xi_0) = \mu$ et (a_n) une suite de réels vérifiant

$$(4.12) \quad \sum_{n \in \mathbb{Z}} a_n^2 < \infty,$$

assurant ainsi la bonne définition du processus. La suite (X_n) est strictement stationnaire et a pour densité spectrale, définie sur le tore,

$$(4.13) \quad f(\theta) = \text{Var}(\xi_0) |g(\theta)|^2 \quad \text{avec} \quad g(\theta) = \sum_{n=-\infty}^{\infty} a_n e^{in\theta}.$$

Notre principale hypothèse sur μ est de supposer qu'elle vérifie une inégalité de Sobolev logarithmique (donc non limitée au cadre gaussien), ou simplement une inégalité de transport T_1 . Nous nous intéresserons ici à l'estimation de la moyenne empirique au niveau des grandes déviations et des déviations modérées, de la covariance empirique (et plus généralement de toutes les fonctionnelles à croissance au plus quadratique) puis finalement à la densité spectrale empirique, au niveau des déviations modérées, complétant ainsi le TCL obtenu par Giraitis-Surgailis [86], les grandes déviations de Bryc-Dembo [40] pour des fonctionnelles quadratiques de processus gaussiens, et les grandes déviations pour les marginales de la densité spectrale empirique de Bercu-Gamboa-Rouault [23].

Soit une suite $1 \ll b_n \ll \sqrt{n}$, et définissons, pour $F : \mathbb{R}^{l+1} \rightarrow \mathbb{R}$,

$$S_n = \sum_{k=1}^n X_k, \quad Q_n^l = \sum_{k=1}^n (X_k X_{k+l} - \mathbb{E}(X_k X_{k+l})),$$

$$S_n(F) = \sum_{k=1}^n (F(X_k, \dots, X_{k+l}) - \mathbb{E}(F(X_k, \dots, X_{k+l}))).$$

Théorème 4.4.1

1. Supposons $\mu \in T_1(C)$, $g \in L^2(d\theta)$, et g continue en 0. Alors, S_n/n vérifie un PGD de vitesse n et de fonction de taux $I(x) = \Lambda^*(X/g(0))$ où Λ^* est la transformée de Legendre de la fonction génératrice des moments de μ . De plus, $S_n/b_n \sqrt{n}$ vérifie un PDM de vitesse b_n^2 et de fonction de taux $I_m(x) = x^2/(2f(0))$.

2. Supposons $\mu \in ISL(C)$, $f \in L^q(d\theta)$ pour $2 < q \leq \infty$ et la condition sur la vitesse $b_n n^{(1-2q)/2q} \rightarrow 0$, alors $(Q_n^l/b_n \sqrt{n})_{0 \leq l \leq m}$ satisfait un PDM sur \mathbb{R}^{m+1} de vitesse b_n^2 et de fonction de taux $J(z) = \sup_{\lambda \in \mathbb{R}^{m+1}} \{ \langle \lambda, z \rangle - \frac{1}{2} \lambda^* \Sigma^2 \lambda \}$ avec $\Sigma^2 = (\Sigma_{k,l}^2)_{0 \leq k, l \leq m}$ et

$$\Sigma_{k,l}^2 = \frac{1}{2\pi} \int \left(e^{i(k-l)\theta} + e^{i(k+l)\theta} \right) f^2(\theta) d\theta + \kappa_4 \left(\frac{1}{2\pi} \int f(\theta) e^{ik\theta} d\theta \right) \left(\frac{1}{2\pi} \int f(\theta) e^{il\theta} d\theta \right).$$

où $\kappa_4 = \frac{\mathbb{E}(\xi^4) - 3\mathbb{E}(\xi^2)^2}{\mathbb{E}(\xi^2)^2}$.

3. Supposons $\mu \in ISL(C)$ et g continue. Supposons de plus que $\delta_{x_i} F$ est Lipschitz pour $i = 0, \dots, l$, alors $S_n(F)/b_n \sqrt{n}$ vérifie un PDM de vitesse b_n^2 et de fonction de taux $I_F(z) = z^2/2\Sigma_F^2$ avec

$$\Sigma_F^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \text{Var} \left(\sum_{k=1}^n F(X_k, \dots, X_{k+l}) \right).$$

Considérons simplement Q_n^l . La preuve repose sur les trois étapes suivantes :

1) on utilise l'approximation de Fejer du processus à moyenne mobile, soit pour $N \in \mathbb{N}^*$, $a_j^N = a_j(1 - |j|/N)$ si $|j| \leq N$ et 0 sinon, et (X_n^N) définie grâce à cette suite est ainsi une suite $2N + l$ -dépendantes. L'inégalité de Sobolev logarithmique et T_1 , par les résultats du chapitre 3, impliquent qu'il existe $\delta > 0$ tel que $\int e^{\delta x^2} d\mu$ est finie et l'on peut donc se ramener aux résultats de déviations modérées de Chen [47] pour les variables m -dépendantes.

2) Il s'agit ici de montrer la négligeabilité asymptotique quand N tend vers l'infini de $\frac{1}{b_n \sqrt{n}}(Q_n^l - Q_n^{l,N})$ soit

$$\lim_{N \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{E} \left(e^{\lambda b_n^2 (Q_n - Q_n^N)} \right) = 0,$$

où nous utilisons une conséquence de l'inégalité de Sobolev logarithmique, *i.e.* pour des fonctions G régulières

$$\mathbb{E} \left(e^{\lambda \frac{b_n}{\sqrt{n}} (G - \mathbb{E}G)} \right) \leq \mathbb{E} \left(e^{\lambda^2 C \frac{b_n^2}{n} |\nabla G|^2} \right).$$

Notant $G((\xi_l)_{l \in \mathbb{Z}}) = \sum_{k=1}^n (X_k X_{k+l} - X_k^N X_{k+l}^N)$, on se réduit donc à des estimations sur le carré du gradient de G qui se décompose en 4 termes pouvant s'écrire chacun de manière générique comme $\langle X_{+l}, T_n((g_g^N)^2) X_{+l} \rangle$, où $T_n(h)$ est la matrice de Toeplitz associée à h . Une procédure de linéarisation originale, basée sur l'intervention d'une mesure gaussienne, permet alors de réduire le problème à l'étude des valeurs propres de

$$\sqrt{T_n((g - g^N)^2)} T_n(f) \sqrt{T_n((g - g^N)^2)},$$

dont le contrôle fait intervenir la condition sur b_n .

3) La négligeabilité démontrée, il reste ensuite à démontrer la convergence des fonctions de taux obtenues en 1), ce qui est réalisé grâce à des estimations issues de 2) et à des inégalités de Hölder.

Ces résultats permettent ensuite de s'intéresser à l'estimateur de la densité spectrale empirique

$$\mathcal{L}_n(\theta) = \frac{\sqrt{n}}{b_n} (\mathcal{I}_n(\theta) - \mathbb{E} \mathcal{I}_n(\theta)).$$

avec $\mathcal{I}_n(\theta) := \frac{1}{n} \left| \sum_{k=1}^n X_k e^{ik\theta} \right|^2$.

Théorème 4.4.2 Supposons $\mu \in LSI(C)$. Supposons de plus que la densité spectrale $f \in L^q(d\theta)$, avec $2 < q \leq +\infty$ et $\frac{b_n}{\sqrt{n}} n^{1/q+1/p'} \rightarrow 0$, $\frac{1}{p} + \frac{1}{p'} = 1$ et $\frac{1}{p'} + \frac{1}{q} < \frac{1}{2}$. Alors $(\mathcal{L}_n)_{n \geq 0}$ satisfait un PDM sur $(L^p(d\theta), \sigma(L^p(d\theta), L^{p'}(d\theta)))$ de vitesse b_n^2 et de fonction de taux donnée, pour $\eta \in L^p(d\theta)$, η paire, par

$$I(\eta) = \begin{cases} \frac{1}{2\pi} \int \frac{\eta^2(\theta)}{4f^2(\theta)} d\theta - \frac{\kappa_4}{2 + \kappa_4} \left(\frac{1}{2\pi} \int \frac{\eta(\theta)}{2f(\theta)} d\theta \right)^2 & \text{si } \eta(\theta) d\theta \ll f(\theta) d\theta \text{ et } \frac{\eta(\theta)}{f(\theta)} \in L^2(d\theta); \\ +\infty, & \text{sinon.} \end{cases}$$

La preuve repose sur le Théorème 4.4.1 et le théorème de limite projective pour lequel il faut montrer

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \left(e^{b_n^2 \frac{1}{2\pi} \int h(\theta) \mathcal{L}_n(\theta) d\theta} \right) < +\infty$$

pour tout $h \in L^{p'}(d\theta)$, ainsi que sur l'identification de la fonction de taux. La preuve de la limite précédente repose encore sur un contrôle de la transformée de Laplace par une utilisation, successivement, de l'inégalité de Sobolev logarithmique puis du procédé de linéarisation, et par le contrôle de valeurs propres associées à un produit de matrices de Toeplitz. L'identification de la fonction de taux se fait par un calcul explicite du maximiseur de la fonction de taux obtenue comme limite projective de celle du théorème 4.4.1.

4.5 Statistique de DURBIN-WATSON

Dans ce travail, en collaboration avec V. BITSEKI PENDA et F. PROÏA [H-D9] (2014), nous nous sommes intéressé au PDM de la statistique de DURBIN-WASTON associée au processus autorégressif de premier ordre stable, pour lequel le bruit est également donné par un processus autorégressif de premier ordre. Pour introduire la statistique de Durbin-Watson, on considère un modèle autorégressif de premier ordre, lui-même issu d'une perturbation autorégressive du premier ordre, définie comme suit : pour tout $n \geq 1$,

$$\begin{cases} X_n &= \theta X_{n-1} + \varepsilon_n \\ \varepsilon_n &= \rho \varepsilon_{n-1} + V_n. \end{cases}$$

Nous supposons ici que les paramètres inconnus vérifient les conditions $|\theta| < 1$ et $|\rho| < 1$ pour assurer la stabilité du modèle. Dans la suite, on suppose que (V_n) est une suite de variables aléatoires indépendantes et identiquement distribuées centrées et de variance positive σ^2 . Les variables aléatoires initiales X_0 et ε_0 peuvent-être choisies d'une manière arbitraire. Nous avons décidé d'estimer θ par l'estimateur des moindres carrés

$$\theta_n = \frac{\sum_{k=1}^n X_k X_{k-1}}{\sum_{k=1}^n X_{k-1}^2}.$$

On définit un ensemble résiduel (ε_k) , sur lequel notre statistique de test pourra s'appuyer, pour tout $1 \leq k \leq n$ soit $\varepsilon_k = X_k - \theta_n X_{k-1}$, avec de manière arbitraire, $\varepsilon_0 = 0$. Cela nous permet d'introduire un estimateur de ρ donné, pour tout $n \geq 1$ par

$$\rho_n = \frac{\sum_{k=1}^n \varepsilon_k \varepsilon_{k-1}}{\sum_{k=1}^n \varepsilon_{k-1}^2}.$$

Finalement, la statistique de Durbin-Watson est définie, pour $n \geq 1$, par

$$D_n = \frac{\sum_{k=1}^n (\varepsilon_k - \varepsilon_{k-1})^2}{\sum_{k=0}^n \varepsilon_k^2}.$$

On sait que si $\mathbb{E}[V_1^2] < \infty$, alors, on a les convergences presque sûres suivantes

$$\lim_{n \rightarrow \infty} \theta_n = \frac{\theta + \rho}{1 + \theta \rho} := \theta^*, \quad \lim_{n \rightarrow \infty} \rho_n = \theta \rho \theta^* := \rho^*, \quad \lim_{n \rightarrow \infty} D_n = 2(1 - \rho^*) := D^*.$$

De plus si $\mathbb{E}[V_1^4] < \infty$ on obtient le TCL :

$$\sqrt{n}(\theta_n - \theta^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\theta^2), \quad \sqrt{n}(\rho_n - \rho^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2) \quad \sqrt{n}(D_n - D^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_D^2),$$

où les variances asymptotiques sont données par

$$\sigma_\theta^2 = \frac{(1 - \theta^2)(1 - \theta\rho)(1 - \rho^2)}{(1 + \theta\rho)^3}, \quad \sigma_D^2 = 4\sigma_\rho^2,$$

$$\sigma_\rho^2 = \frac{(1 - \theta\rho)}{(1 + \theta\rho)^3} ((\theta + \rho)^2(1 + \theta\rho)^2 + (\theta\rho)^2(1 - \theta^2)(1 - \rho^2)).$$

Notre objectif est maintenant d'établir le PDM pour ces estimateurs, afin d'obtenir une meilleure asymptotique que le TCL. Dans ce premier résultat, nous nous concentrons sur le PDM de la statistique de Durbin-Watson dans le cas simple où le bruit (V_n) est Gaussien. Cette condition restrictive nous permet de réduire les hypothèses à l'existence d'un $t > 0$ tel que

$$(\mathbf{G}) \quad \mathbb{E}[\exp(t\varepsilon_0^2)] < \infty \quad \text{et} \quad \mathbb{E}[\exp(tX_0^2)] < \infty.$$

Théorème 4.5.1 *Supposons la condition (\mathbf{G}) satisfaite et que (V_n) suit la loi $\mathcal{N}(0, \sigma^2)$. Alors*

1. *à condition que $\theta \neq -\rho$, la suite*

$$\left(\frac{\sqrt{n}}{b_n} \begin{pmatrix} \theta_n - \theta^* \\ \rho_n - \rho^* \end{pmatrix} \right)_{n \geq 1}$$

satisfait un PGD sur \mathbb{R}^2 de vitesse b_n^2 et de fonction de taux $K(x) = \frac{1}{2}x'\Gamma^{-1}x$ où Γ est donnée par

$$\Gamma = \begin{pmatrix} \sigma_\theta^2 & \theta\rho\sigma_\theta^2 \\ \theta\rho\sigma_\theta^2 & \sigma_\rho^2 \end{pmatrix}.$$

2. *la suite $(\sqrt{n}(D_n - D^*)/b_n)_{n \geq 1}$ satisfait un PGD sur \mathbb{R} de vitesse b_n^2 et de fonction de taux $I_D(x) = x^2/2\sigma_D^2$.*

En outre, dans le cas particulier où $\theta = -\rho$, les suites $(\sqrt{n}(\theta_n - \theta^*)/b_n)_{n \geq 1}$ et $(\sqrt{n}(\rho_n - \rho^*)/b_n)_{n \geq 1}$ satisfont un PGD sur \mathbb{R} de vitesse b_n^2 et de fonction de taux données respectivement par

$$I_\theta(x) = \frac{x^2(1 - \theta^2)}{2(1 + \theta^2)} \quad \text{et} \quad I_\rho(x) = \frac{x^2(1 - \theta^2)}{2\theta^4(1 + \theta^2)}.$$

Nous concentrons notre attention maintenant sur le cadre plus général où le bruit (V_n) est supposé satisfaire une condition de type de Chen-Ledoux.

Théorème 4.5.2 *Sous les hypothèses suivantes*

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log n \mathbb{P}(|V_1|^a > b_n \sqrt{n}) = -\infty \quad \frac{|\varepsilon_0|^4}{b_n \sqrt{n}} \xrightarrow{\text{superexp}}_{b_n^2} 0, \quad \frac{|X_0|^4}{b_n \sqrt{n}} \xrightarrow{\text{superexp}}_{b_n^2} 0,$$

les conclusions du Théorème 4.5.1 restent vraies.

Pour la preuve, nous nous sommes beaucoup inspirés des travaux récents de BERCU et PROÏA dans lesquels la convergence presque sûre et le TCL ont été établis. Nos résultats sont prouvés en utilisant des résultats de Dembo [61], Dembo et Zeitouni [64] et Worms [170] d'une part et des articles de Puhalskii [147] et Djellout [T-D4] d'autre part, sur le PDM des martingales.

Perspectives et problèmes

Les résultats précédents se placent clairement dans le cadre des processus à mémoire courte. Leur extension à des processus à moyenne mobile, à mémoire longue est un domaine complètement vierge tant au niveau des déviations modérées qu'au niveau des grandes déviations. Outre l'intérêt pratique évident de considérer des processus à mémoire longue (présents en finance, télécommunications ...), un aspect théorique intéressant se dégage : les théorèmes du second ordre obtenus montrent des convergence vers des limites non gaussiennes et il serait alors très intéressant d'étudier la fonction de taux des déviations modérées, qui ne sera vraisemblablement pas quadratique.

Bibliographie

- [1] Y. Aït-Sahalia. Disentangling diffusion from jumps. *Journal of Financial Economics*, 74 :487–528, 2004.
- [2] Y. Aït-Sahalia, J. Fan, and D. Xui. High-frequency covariance estimates with noisy and asynchronous financial data. *J. Amer. Statist. Assoc.*, 105(492) :1504–1517, 2010.
- [3] Y. Aït-Sahalia and J. Jacod. Estimating the degree of activity of jumps in high frequency data. *The Annals of Statistics*, 37(5A) :2202–2244, 2009.
- [4] Y. Aït-Sahalia and J. Jacod. Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1) :184–222, 2009.
- [5] T. G. Andersen, T. Bollerslev, T. X. Diebold, and G. Wu. Realized beta : persistence and predictability. *Adv. Econom.*, 20 :1–39, 2008.
- [6] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Schefker. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses*. Société Mathématique de France, Paris, 2000.
- [7] M. A. Arcones. The large deviation principle for stochastic processes I, II. *Theor. Probab. Appl.*, 47(4) : 567–583 :48(1) : 19–44, 2003.
- [8] M. A. Arcones. Moderate deviations of empirical processes. In *Stochastic inequalities and applications*, volume 56 of *Progr. Probab.*, pages 189–212. Birkhäuser, Basel, 2003.
- [9] K. B. Athreya and P. Ney. A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.*, 245 :493–501, 1978.
- [10] R. Azencott. Grandes déviations et applications. In *Eighth Saint Flour Probability Summer School—1978*, volume 774 of *Lecture Notes in Math.*, pages 1–176. Springer, Berlin, 1980.
- [11] K. Azuma. Weighted sums of certain dependent random variables. *Tôhoku Math. J. (2)*, 19 :357–367, 1967.
- [12] P. Baldi. Large deviations and stochastic homogenization. *Ann. Mat. Pura Appl.*, 151 :161–177, 1988.
- [13] V. Bansaye, J. F. Delmas, L. Marsalle, and V. C. Tran. Limit theorems for Markov processes indexed by continuous time Galton-Watson trees. *Ann. Appl. Probab.*, 21(6) :2263–2314, 2011.
- [14] O. E. Barndorff-Nielsen, S.E. Graversen, J. Jacod, M. Podolskij, and N. Shephard. A central limit theorem for realised power and bipower variations of continuous semimartingales. *From stochastic calculus to mathematical finance*, pages 33–68, 2006.
- [15] O. E. Barndorff-Nielsen, S.E. Graversen, J. Jacod, and N. Shephard. Limit theorems for bipower variation in financial econometrics. *Econometric Theory*, 22(4) :677–719, 2006.
- [16] O. E. Barndorff-Nielsen and N. Shephard. Econometric analysis of realized volatility and its use in estimating volatility models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 64(2) :253–280, 2002.
- [17] O. E. Barndorff-Nielsen, N. Shephard, and M. Winkel. Limit theorems for multipower variation in the presence of jumps. *Stochastic Process. Appl.*, 116 :796–806, 2006.
- [18] I. V. Basawa and J. Zhou. Non-Gaussian bifurcating models and quasi-likelihood estimation. *J. Appl. Probab.*, 41A :55–64, 2004.

- [19] G. Bennett. Probability inequalities for sum of independent random variables. *J. Amer. Statist. Assoc.*, 57(297) :33–45, Mar. 1962.
- [20] B. Bercu and V. Blandin. Limit theorems for bifurcating integer-valued autoregressive processes. *Stat. Inference Stoch. Process.*, 18(1) :33–67, 2015.
- [21] B. Bercu and V. Blandin. A Rademacher-Menchov approach for random coefficient bifurcating autoregressive processes. *Stochastic Process. Appl.*, 125(4) :1218–1243, 2015.
- [22] B. Bercu, B. de Saporta, and A. Gégout-Petit. Asymptotic analysis for bifurcating autoregressive processes via a martingale approach. *Electron. J. Probab.*, 14 :no. 87, 2492–2526, 2009.
- [23] B. Bercu, F. Gamboa, and A. Rouault. Large deviations for quadratic forms of stationary Gaussian processes. *Stochastic Process. Appl.*, 71(1) :75–90, 1997.
- [24] B. Bercu and F. Proïa. A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first-order autoregressive process. *ESAIM Probab. Stat.*, 16, 2012.
- [25] B. Bercu and A. Touati. Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.*, 18(5), 2008.
- [26] M. Bibinger. An estimator for the quadratic covariation of asynchronously observed Itô processes with noise : asymptotic distribution theory. *Stochastic Process. Appl.*, 122(6) :2411–2453, 2012.
- [27] M. Bibinger and M. Reiß. Spectral estimation of covolatility from noisy observations using local weights. *Scand. J. Stat.*, 41(1) :23–50, 2014.
- [28] V. Blandin. Asymptotic results for random coefficient bifurcating autoregressive processes. *Statistics*, 48(6) :1202–1232, 2014.
- [29] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1) :1–28, 1999.
- [30] S. G. Bobkov and C. Houdre. Isoperimetric constants for product probability measures. *Ann. Probab.*, 25(1) :184–205, 1997.
- [31] S. G. Bobkov and M. Ledoux. Weighted Poincaré-type inequalities for Cauchy and other convex measures. *Ann., Probab.*, 37(2) :403–427, 2009.
- [32] S.G. Bobkov, I. Gentil, and M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pu. Appl.*, 80(7) :669–696, 2001.
- [33] S.G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1) :1–28, 1999.
- [34] S.G. Bobkov and M. Ledoux. Poincaré’s inequalities and Talagrand’s concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3) :383–400, 1997.
- [35] S.G. Bobkov and M. Ledoux. From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. *Geom. Funct. Anal.*, 10(5) :1028–1052, 2000.
- [36] F. Bolley and C. Villani. Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities. *Ann. Fac. Sci. Toulouse Math. (6)*, 14(3) :331–352, 2005.
- [37] A. A. Borovkov. Boundary value problems for random walks and large deviations in function spaces. *Theor. Probability Appl.*, 12 :575–595, 1967.
- [38] A. A. Borovkov and A. A. Mogulskii. Probabilities of large deviations in topological vector space I, II. *Siberian Maths J.*, 19-21 :161–177, 12–26, 1978-1980.
- [39] W. Bryc and A. Dembo. On large deviations of empirical measures for stationary Gaussian processes. *Stochastic Process. Appl.*, 58(1) :23–34, 1995.
- [40] W. Bryc and A. Dembo. Large deviations for quadratic functionals of Gaussian processes. *J. Theoret. Probab.*, 10(2) :307–332, 1997. Dedicated to Murray Rosenblatt.
- [41] P. Buser. A note on the isoperimetric constant. *Ann. Sci. École Norm. Sup*, 15(4) :213–230, 1982.
- [42] M. Capitaine, E.P. Hsu, and M. Ledoux. Martingale representation and a simple proof of logarithmic sobolev inequality on path spaces. *Elect. Comm. Probab.*, 2 :71–81, 1997.

- [43] M.F. Chen. Analytic proof of dual variational formula for the first eigenvalue in dimension one. *Sci. China Ser. A*, 42(8) :805–815, 1999.
- [44] M.F. Chen. *Eigenvalues, inequalities, and ergodic theory*. Probability and its Applications. Springer-Verlag, 2005.
- [45] M.F. Chen and F. Y. Wang. Estimation of the first eigenvalue of second order elliptic operators. *J. Funct. Anal.*, 131(2) :345–363, 1995.
- [46] X. Chen. Moderate deviations for m -dependent random variables with Banach space values. *Statist. Probab. Lett.*, 35(2) :123–134, 1997.
- [47] X. Chen. Limit theorems for functionals of ergodic Markov chains with general state space. *Mem. Amer. Math. Soc.*, 139(664) :xiv+203, 1999.
- [48] X. Chen. On the limit laws of the second order for additive functionals of Harris recurrent Markov chains. *Probab. Theory Related Fields*, 116(1) :89–123, 2000.
- [49] F. Comte, V. Genon-Catalot, and Y. Rozenholc. Nonparametric estimation for a stochastic volatility model. *Finance Stoch.*, 14(1) :49–80, 2010.
- [50] R. Cowan and R. G. Staudte. The bifurcating autoregressive model in cell lineage studies. *Biometrics*, 42(4) :769–783, 1986.
- [51] A. Dalalyan and N. Yoshida. Second-order asymptotic expansion for a non-synchronous covariation estimator. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(3) :748–789, 2011.
- [52] A. de Acosta. Large deviations for vector-valued Lévy processes. *Stochastic Process. Appl.*, 51(1) :75–115, 1994.
- [53] A. de Acosta. Moderate deviations for empirical measures of Markov chains : lower bounds. *Ann. Probab.*, 25(1) :259–284, 1997.
- [54] A. de Acosta and X. Chen. Moderate deviations for empirical measures of Markov chains : upper bounds. *J. Theoret. Probab.*, 11(4) :1075–1110, 1998.
- [55] V. H. de la Peña, T. L. Lai, and Q. M. Shao. *Self-normalized processes*. Probability and its Applications (New York). Springer-Verlag, Berlin, 2009.
- [56] B. de Saporta, A. Gégout-Petit, and L. Marsalle. Parameters estimation for asymmetric bifurcating autoregressive processes with missing data. *Electron. J. Stat.*, 5 :1313–1353, 2011.
- [57] B. de Saporta, A. Gégout-Petit, and L. Marsalle. Asymmetry tests for bifurcating auto-regressive processes with missing data. *Statist. Probab. Lett.*, 82(7) :1439–1444, 2012.
- [58] B. de Saporta, A. Gégout-Petit, and L. Marsalle. Random coefficients bifurcating autoregressive processes. *ESAIM Probab. Stat.*, 18 :365–399, 2014.
- [59] S. Delattre and J. Jacod. A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors. *Bernoulli*, 3(1) :1–28, 1997.
- [60] J.F. Delmas and L. Marsalle. Detection of cellular aging in a Galton-Watson process. *Stochastic Process. Appl.*, 120(12) :2495–2519, 2010.
- [61] A. Dembo. Moderate deviations for martingales with bounded jumps. *Electron. Comm. Probab.*, 1, no. 3 :11–17, 1996.
- [62] A. Dembo. Information inequalities and concentration of measure. *Ann. Probab.*, 25(1) :927–939, 1997.
- [63] A. Dembo and T. Zajic. Uniform large and moderate deviations for functional empirical processes. *Stochastic Process. Appl.*, 67(2) :195–211, 1997.
- [64] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.
- [65] N. Demni and M. Zani. Large deviations for statistics of the jacobi process. *Stochastic Process. Appl.*, 119(2) :518–533, 2009.
- [66] J.D. Deuschel and D.W. Stroock. *Large deviations*, volume 137 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1989.

- [67] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I. II. *Comm. Pure Appl. Math.*, 28, 1975.
- [68] M. D. Donsker and S. R. S. Varadhan. Large deviations for stationary Gaussian processes. *Comm. Math. Phys.*, 97(1-2) :187–210, 1985.
- [69] P. Dovonon, S. Gonçalves, and N. Meddahi. Bootstrapping realized multivariate volatility measures. *J. Econometrics*, 172(1) :49–65, 2013.
- [70] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23(4) :1671–1691, 1995.
- [71] P. Dupuis and R.S. Ellis. *A weak convergence approach to the theory of large deviations*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., New York, 1997.
- [72] J. Durbin. Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables. *Econometrica*, 38 :410–421, 1970.
- [73] J. Durbin and G. S. Watson. Testing for serial correlation in least squares regression. I, II, III. *Biometrika*, 37, 38,58 :409–428, 159–178, 1–19, 1950,1951,1971.
- [74] A. Eberle. *Uniqueness and Non-Uniqueness of Semigroups Generated by Singular Diffusion Operators*. Lecture Notes in Mathematics, 1718. Springer-Verlag, Berlin, 1999.
- [75] P. Eichelsbacher and M. Löwe. Moderate deviations for i.i.d. random variables. *ESAIM Probab. Stat.*, 7 :209–218, 2003.
- [76] R.S. Ellis. *Entropy, large deviations, and statistical mechanics*, volume 271 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, New York, 1985.
- [77] F. F. Barthe and C Roberto. Sobolev inequalities for probability measures on the real line. *Studia Math.*, 159(3) :481–497, 2003.
- [78] J. E. Figueroa-López and J. Nisen. Optimally thresholded realized power variations for Lévy jump diffusion models. *Stoch. Proc. Appl.*
- [79] D. Florens-Zmirou. On estimating the diffusion coefficient from discrete observations. *J. Appl. Probab.*, 30(4) :790–804, 1993.
- [80] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, New York, 1984.
- [81] M. Fukasawa. Realized volatility with stochastic sampling. *Stochastic Process. Appl.*, 120(6) :829–852, 2010.
- [82] F. Gamboa, A. Rouault, and M. Zani. A functional large deviations principle for quadratic forms of Gaussian stationary processes. *Statist. Probab. Lett.*, 43(3) :299–308, 1999.
- [83] F. Q. Gao. Uniform moderate deviations of Markov processes-the discrete parameter case. *Acta Math. Sinica*, 38(4) :543–552, 1995.
- [84] F. Q. Gao. Moderate deviations for martingales and mixing random processes. *Stoch. Proc. Appl.*, 61 :263–275, 1996.
- [85] F. Q. Gao and X. Zhao. Delta method in large deviations and moderate deviations for estimators. *Ann. Statist.*, 39(2) :1211–1240, 2011.
- [86] L. Giraitis and D. Surgailis. A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle’s estimate. *Probab. Theory Related Fields*, 86(1), 1990.
- [87] A. Gloter. Efficient estimation of drift parameters in stochastic volatility models. *Finance Stoch.*, 11(4) :495–519, 2007.
- [88] F. Gobbi and C. Mancini. Estimating the diffusion part of the covariation between two volatility models with jumps of Lévy type. 75 :399–409, 2007.
- [89] M. I. Gordin. The central limit theorem for stationary processes. *Soviet Math. Dokl.*, 10 :1174–1176, 1969.
- [90] N. Gozlan. Integral criteria for transportation-cost inequalities. *Electron. Comm. Probab.*, 11 :64–77 (electronic), 2006.

- [91] N. Gozlan and C. Léonard. A large deviation approach to some transportation cost inequalities. *Probab. Theory Related Fields*, 139(1-2) :235–283, 2007.
- [92] I. Grama and E. Haeusler. Large deviations for martingales via cramèr’s method. *Stoch. Proc. Appl.*, 85 :279–260, 2000.
- [93] A. Guillin. Uniform moderate deviations of functional empirical processes of markov chains. *Probab. Math. Statist.*, 20 :237–260, 2000.
- [94] A. Guillin, C. Léonard, F. Y. Wang, and L. Wu. Transportation-information inequalities for Markov processes (II) : relations with other functional inequalities. *J. Geom. Anal.*, 18(4) :921–979, 2008.
- [95] A. Guillin, C. Léonard, L. Wu, and N. Yao. Transport-information inequalities for Markov processes (I). *Probab. Theory Related Fields*, 144(3-4) :669–695, 2009.
- [96] J. Guyon. Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.*, 17(5-6) :1538–1569, 2007.
- [97] J. Guyon, A. Bize, G. Paul, E. Stewart, J. F. Delmas, and F. Taddéi. Statistical study of cellular aging. In *CEMRACS 2004—mathematics and applications to biology and medicine*, volume 14 of *ESAIM Proc.*, pages 100–114 (electronic). EDP Sci., Les Ulis, 2005.
- [98] P. Hall and C. C. Heyde. *Martingale Limit Theory and its Application*. Academic Press, 1980.
- [99] T. Hayashi and N. Yoshida. Nonsynchronous covariation process and limit theorems. *Stochastic Process. Appl.*, 121(10) :2416–2454, 2011.
- [100] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58 :13–30, 1963.
- [101] R. M. Huggins and I. V. Basawa. Extensions of the bifurcating autoregressive model for cell lineage studies. *J. Appl. Probab.*, 36(4) :1225–1233, 1999.
- [102] R. M. Huggins and I. V. Basawa. Inference for the extended bifurcating autoregressive model for cell lineage studies. *Aust. N. Z. J. Stat.*, 42(4) :423–432, 2000.
- [103] S. Y. Hwang, I. V. Basawa, and I. K. Yeo. Local asymptotic normality for bifurcating autoregressive processes and related asymptotic inference. *Stat. Methodol.*, 6(1) :61–69, 2009.
- [104] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. Second Edition North-Holland Mathematical Library, 1989.
- [105] J. Jacod, Y.Y. Li, P.A. Mykland, M. Podolskij, and M. Vetter. Microstructure noise in the continuous case : the pre-averaging approach. *Stochastic Process. Appl.*, 119(7) :2249–2276, 2009.
- [106] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, 1987.
- [107] H. Jiang. Moderate deviations for estimators of quadratic variational process of diffusion with compound poisson. *Statist. Probab. Lett.*, 80(17-18) :1297–1305, 2010.
- [108] S. Kanaya and T. Otsu. Large deviations of realized volatility. *Stochastic Process. Appl.*, 122(2) :546–581, 2012.
- [109] S. Kanaya and T. Otsu. Corrigendum to “Large deviations of realized volatility” [Stochastic Process. Appl. 122 (2012) 546–581]. *Stochastic Process. Appl.*, 123(3) :1176–1177, 2013.
- [110] M. L. King and P. X. Wu. Small-disturbance asymptotics and the Durbin-Watson and related tests in the dynamic regression model. *J. Econometrics*, 47 :145–152, 1991.
- [111] T. Klein, Y. Ma, and N. Privault. Convex concentration inequality and forward/backward martingale stochastic calculs. *Electron. J. Probab.*, 11(20) :486–512, 2006.
- [112] M. Ledoux. Measure concentration, transportation cost, and functional inequalities. *Instructional Conference on Combinatorial Aspects of Mathematical Analysis, Edinburgh, 25 March-5 April 2002 and Summer School on Singular Phenomena and Scaling in Mathematical Models, Bonn, 10-13 June 2003*.
- [113] M. Ledoux. Sur les déviations modérées des sommes de variables aléatoires vectorielles indépendantes de même loi. *Ann. Inst. H. Poincaré*, 28 :267–280, 1992.
- [114] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. *Séminaire de Probabilités, XXXIII, Lecture Notes in Math. Springer, Berlin*, (1709) :120–216, 1999.

- [115] M. Ledoux. *The concentration of measure phenomenon*. Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI, 2001.
- [116] M. Ledoux. Spectral gap, logarithmic sobolev constant, and geometric bounds. *Surv. Differ. Geom.*, Somerville, IX :219–240, 2004.
- [117] R. S. Liptser and A. A. Pukhalskii. Limit theorems on large deviations for semimartingales. *Stochastics Stochastics Rep.*, 38(4) :201–249, 1992.
- [118] O. Ludger. Estimation for continuous branching processes. *Scand. J. Statist.*, 25(1) :111–126, 1998.
- [119] J. Lynch and J. Sethuraman. Large deviations for processes with independent increments. *Ann. Probab.*, 15(2) :610–627, 1987.
- [120] M. Maïda, J. Najim, and S. Péché. Large deviations for weighted empirical mean with outliers. *Stochastic Process. Appl.*, 117(10) :1373–1403, 2007.
- [121] C. Mancini. Estimation of the characteristics of the jumps of a general Poisson-diffusion model. *Scand. Actuar. J.*, (1) :42–52, 2004.
- [122] C. Mancini. Large deviation principle for an estimator of the diffusion coefficient in a jump-diffusion process. *Statist. Probab. Lett.*, 78(7) :869–879, 2008.
- [123] C. Mancini and F. Gobbi. Identifying the brownian covariation from the co-jumps given discrete observations. *Econometric Theory*, 28(2) :249–273, 2012.
- [124] K. Marton. Bounding \bar{d} -distance by information divergence : a method to prove measure concentration. *Ann. Probab.*, 24(2) :857–866, 1996.
- [125] K. Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6(3) :556–571, 1996.
- [126] K. Marton. Measure concentration for a class of random processes. *Probab. Theory Relat. Fields*, 110(3) :427–439, 1998.
- [127] P. Massart. *Concentration inequalities and model selection*, volume 1896 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.
- [128] C. McDiarmid. *On the method of bounded differences*. Surveys of Combinatorics, J. Siemons (Ed.), London Math. Soc. Lect. Notes Series, 1989.
- [129] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Communications and Control Engineering. Springer-Verlag London Ltd., London, 1993.
- [130] E. Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Invent. Math.*, 177(1) :1–43, 2009.
- [131] A. A. Mogul'skiĭ. Large deviations for the trajectories of multidimensional random walks. *Theor. Probability Appl.*, 21(2) :300–315, 1976.
- [132] A. A. Mogul'skiĭ. Large deviations for processes with independent increments. *Ann. Probab.*, 21(1) :202–215, 1993.
- [133] J. Najim. Grandes déviations pour certaines mesures empiriques. *PhD thesis, Université Nanterre Paris-10*, pages 1–130, 2001.
- [134] J. Najim. A Cramér type theorem for weighted random variables. *Electron. J. Probab.*, 7(4) :1–32, 2002.
- [135] J. Najim. Large deviations for independent random variables - Application to Erdős-Rényi's functional law of large numbers. *ESAIM Probab. Stat.*, 9 :116–142 (electronic), 2005.
- [136] E. Nummelin. A splitting technique for Harris recurrent Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 4 :309–318, 1978.
- [137] E. Nummelin. *General irreducible Markov chains and nonnegative operators*. Cambridge tract in mathematics. Cambridge University Press, 1984.
- [138] E. Nummelin and P. Tuominen. Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. *Stoch. Proc. Appl.*, 12 :187–202, 1982.

- [139] E. Nummelin and R. L. Tweedie. Geometric ergodicity and R-positivity for general Markov chains. *Ann. Prob.*, 6 :404–420, 1978.
- [140] T. Ogihara and N. Yoshida. Quasi-likelihood analysis for the stochastic differential equation with jumps. *Stat. Inference Stoch. Process.*, 14(3) :189–229, 2011.
- [141] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2) :361–400, 2000.
- [142] O. Perrin and M. Zani. Large deviations for sample paths of Gaussian processes quadratic variations. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 328(Veroyatn. i Stat. 9) :169–181, 280, 2005.
- [143] M. Podolskij and M. Vetter. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3) :634–658, 2009.
- [144] F. Proïa. Further results on the h -test of Durbin for stable autoregressive processes. *J. Multivariate Anal.*, 118 :77–101, 2013.
- [145] A. Puhalskii. Large deviations of semimartingales via convergence of the predictable characteristics. *Stoch. Stoch. Rep.*, 49(1-2) :27–85, 1994.
- [146] A. Puhalskii. The method of stochastic exponentials for large deviations. *Stoch. Proc. Appl.*, 54 :45–70, 1994.
- [147] A. Puhalskii. Large deviations of semimartingales : a maxingale problem approach. I. Limits as solutions to a maxingale problem. II. Uniqueness of the maxingale problem. Application. *Stoch. Stoch. Rep.*, 61-68 :141–243 :65–143, 1997-1999.
- [148] A. Račkauskas. On probabilities of large deviations for martingales. *Leituvos Matematikos Rinkiny*s, pages 784–794, 1990.
- [149] A. Račkauskas. Large deviations for martingales with some applications. *Acta Applicandae Mathematicae*, 38 :109–129, 1995.
- [150] C. Revuz and M. Yor. *Continuous martingales and Brownian motion. Third edition.* Springer-Verlag, Berlin, 1999.
- [151] E. Rio. Inégalités de Hoeffding pour les fonctions Lipschitziennes de suites dépendantes. *CRAS, Paris*, T.330(Serie I) :905–908, 2000.
- [152] C. Y. Robert and M. Rosenbaum. Volatility and covariation estimation when microstructure noise and trading times are endogenous. *Math. Finance*, 22(1) :133–164, 2012.
- [153] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
- [154] P.M. Samson. Concentration of measure inequalities for Markov chains and ϕ -mixing process. *Ann. Probab.*, 28(1) :416–461, 2000.
- [155] B. Saussereau. Transportation inequalities for stochastic differential equations driven by a fractional Brownian motion. *Bernoulli*, 18(1) :1–23, 2012.
- [156] A. Schied. Cramér’s condition and Sanov’s theorem. *Statist. Probab. Lett.*, 39, 1998.
- [157] Y. Shimizu. Functional estimation for Lévy measures of semimartingales with Poissonian jumps. *J. Multivariate Anal.*, 100(6) :1073–1092, 2009.
- [158] Y. Shimizu. Threshold estimation for jump-type stochastic processes from discrete observations. *Proceedings of the Institute of Statistical Mathematics*, 57(1) :97–118, 2009.
- [159] Y. Shimizu and N. Yoshida. Estimation of parameters for diffusion processes with jumps from discrete observations. *Stat. Inference Stoch. Process.*, 9(3) :227–277, 2006.
- [160] E. J. Stewart, R. Madden, G. Paul, and F. Taddéi. Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol*, 3(2) :e45, 2005.
- [161] T. Stocker. On the asymptotic bias of OLS in dynamic regression model with autocorrelated errors. *Statist. Papers*, 48 :81–93, 2007.
- [162] W. F. Stout. A martingale analogue of Kolmogorov’s law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 15 :279–290, 1970.

- [163] M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3) :587–600, 1996.
- [164] V. Todorov and T. Bollerslev. Jumps and betas : a new framework for disentangling and estimating systematic risks. *J. Econometrics*, 157(2) :220–235, 2010.
- [165] A. W. Van Der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.
- [166] S. R. S. Varadhan. Asymptotic probabilities and differential equations. *Comm. Pure Appl. Math.*, 19 :261–286, 1966.
- [167] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [168] F. Y. Wang. *Functional inequalities, Markov Semigroup and Spectral Theory*. Chinese Sciences Press, Beijing/New York, 2005.
- [169] J. Worms. Moderate deviations for stable Markov chains and regression models. *Electron. J. Probab.*, 4, no. 8 :1–28, 1999.
- [170] J. Worms. Moderate deviations of some dependent variables. I. Martingales. II. Some kernel estimators. *Math. Methods Statist.*, 10(1-2), 2001.
- [171] L. Wu. Large deviations, moderate deviations and LIL for empirical processes. *Ann. Probab.*, 22(1) :17–27, 1994.
- [172] L. Wu. Moderate deviations of dependent random variables related to CLT. *Ann. Probab.*, 23(1) :420–445, 1995.
- [173] L. Wu. Exponential convergence in probability for empirical means of Brownian motion and of random walks. *J. Theoret. Probab.*, 12(3) :661–673, 1999.
- [174] L. Wu. A deviation inequality for non-reversible Markov processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(4) :435–445, 2000.
- [175] L. Wu. Uniformly integrable operators and large deviations for Markov processes. *J. Funct. Anal.*, 172(2) :301–376, 2000.
- [176] L. Wu. Uniqueness of Nelson’s diffusions. II. Infinite dimensional setting and applications. *Potential Anal.*, 13(3) :269–301, 2000.
- [177] L. Wu. Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Process. Appl.*, 91(2) :205–23, 2001.
- [178] L. Wu. Essential spectral radius for Markov semigroups. I. Discrete time case. *Probab. Theory Related Fields*, 128(2) :255–321, 2004.
- [179] L. Wu. Gradient estimates of Poisson equations on Riemannian manifolds and applications. *J. Funct. Anal.*, 257(12) :4015–4033, 2009.
- [180] C. Yu, Y. Fang, Z. Li, B. Zhang, and X. Zhao. Non-parametric estimation of high-frequency spot volatility for Brownian semimartingale with jumps. *J. Time Series Anal.*, 35(6) :572–591, 2014.
- [181] M. Zani. Grandes déviations pour des fonctionnelles issues de la statistique des processus. *PhD thesis, Université Paris-Sud*, (4) :1–151, 1999.
- [182] M. Zani. Sample path large deviations for squares of stationary Gaussian processes. *Theory Probab. Appl.*, 57(2) :347–357, 2013.
- [183] L. Zhang. Efficient estimation of stochastic volatility using noisy observations : a multi-scale approach. *Bernoulli*, 12(6) :1019–1043, 2006.
- [184] L. Zhang, P. M. Mykland, and Y. Aït-Sahalia. A tale of two time scales : determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.*, 100(472) :1394–1411, 2005.
- [185] L. Zhang, P. M. Mykland, and Y. Aït-Sahalia. Edgeworth expansions for realized volatility and related estimators. *J. Econometrics*, 160 :190–203, 2011.
- [186] J. Zhou and I. V. Basawa. Least-squares estimation for bifurcating autoregressive processes. *Statist. Probab. Lett.*, 74(1) :77–88, 2005.
- [187] J. Zhou and I. V. Basawa. Maximum likelihood estimation for a first-order bifurcating autoregressive process with exponential errors. *J. Time Ser. Anal.*, 26(6) :825–842, 2005.